

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 4, 2003, 04:39:33 ; Search time 188 Seconds  
(without alignments)  
8439.558 Million cell updates/sec

Title: US-10-010-160-1

Perfect score: 622

Sequence: 1 atgtctgatgatccagctaa.....aattatgcagtcgacagtc 622

Scoring table: IDENTITY NUC

Searched: 1708419 seqs, 1275431651 residues

Total number of hits satisfying chosen parameters: 3416838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	622	100.0	622	US-10-010-160-1	Sequence 1, Appl
2	54.8	8.8	9052	US-10-311-455-118	Sequence 118, App
3	53	8.5	945	US-09-895-913A-45	Sequence 45, Appl
4	52.8	8.5	7167	US-10-311-455-374	Sequence 374, App
5	52	8.4	7128	US-10-311-455-1532	Sequence 1532, App
6	51.8	8.3	5523	US-10-311-455-137	Sequence 137, App
7	51.8	8.3	73334	US-10-311-455-2097	Sequence 2097, App
8	51.2	8.2	12025	US-10-311-455-1272	Sequence 1272, App
9	51	8.2	6118	US-10-311-455-1004	Sequence 1004, App
10	51	8.2	7921	US-10-311-455-1944	Sequence 1944, App
11	50.6	8.1	6126	US-10-311-455-1804	Sequence 1804, App
12	50.6	8.1	11047	US-10-311-455-1958	Sequence 1958, App
13	50.6	8.1	11047	US-10-240-453-276	Sequence 276, App
14	50.6	8.1	11047	US-10-239-676-188	Sequence 188, App
15	50.2	8.1	51509	US-09-754-853A-4	Sequence 4, Appl
16	49.8	8.0	5935	US-10-239-676-133	Sequence 133, App

17	49.8	8.0	40862	US-10-311-455-2045	Sequence 2045, App
18	48.8	7.8	5542	US-10-311-455-1994	Sequence 1994, App
19	48.6	7.8	18855	US-10-311-455-583	Sequence 583, App
20	48.2	7.7	12138	US-10-311-455-1602	Sequence 1602, App
21	48.2	7.7	12590	US-10-311-455-1893	Sequence 1893, App
22	48.2	7.7	18512	US-10-311-455-950	Sequence 950, App
23	48	7.7	13427	US-10-311-455-1900	Sequence 1900, App
24	47.8	7.7	19972	US-10-311-455-180	Sequence 80, Appl
25	47.6	7.7	7110	US-10-311-455-1530	Sequence 1530, App
26	47.6	7.7	11092	US-10-311-455-1486	Sequence 1486, App
27	47.4	7.6	6056	US-10-311-455-999	Sequence 999, App
28	47.4	7.6	6059	US-10-311-455-1454	Sequence 1454, App
29	47.4	7.6	6385	US-10-240-453-57	Sequence 57, Appl
30	47.4	7.6	6385	US-10-239-676-57	Sequence 57, Appl
31	47.4	7.6	7057	US-10-311-455-1892	Sequence 1822, App
32	47.4	7.6	7057	US-10-240-485-148	Sequence 148, App
33	47.2	7.6	17721	US-10-311-455-1702	Sequence 1702, App
34	47.2	7.6	19787	US-10-311-455-1424	Sequence 1424, App
35	47	7.6	11416	US-10-311-455-92	Sequence 92, Appl
36	46.8	7.5	11805	US-10-311-455-1722	Sequence 1722, App
37	46.8	7.5	15743	US-10-240-453-269	Sequence 269, App
38	46.6	7.5	6254	US-10-311-455-1593	Sequence 1593, App
39	46.6	7.5	6317	US-10-204-708-11	Sequence 11, Appl
40	46.6	7.5	6317	US-10-311-455-381	Sequence 381, App
41	46.6	7.5	11812	US-10-311-455-2092	Sequence 2092, App
42	46.6	7.5	11812	US-10-240-453-306	Sequence 306, App
43	46.6	7.5	11812	US-10-239-676-210	Sequence 210, App
44	46.4	7.5	6089	US-10-311-455-1810	Sequence 1810, App
45	46.2	7.4	6073	US-10-311-455-1515	Sequence 1515, App

## ALIGNMENTS

RESULT 1  
US-10-010-160-1  
Sequence 1, Application US/10010160  
Publication No. US20030103999A1  
GENERAL INFORMATION:  
APPLICANT: Roesey, Everett L.  
APPLICANT: Strugnell, Richard A.  
APPLICANT: Good, Robert T,  
APPLICANT: King, Kendall W.  
TITLE OF INVENTION: NOVEL THERAPEUTIC COMPOSITIONS FOR  
TREATING INFECTION BY LAWSONIA SPP.  
FILE REFERENCE: DAV1110.001AUS  
CURRENT APPLICATION NUMBER: US/10/010,160  
CURRENT FILING DATE: 2001-11-09  
PRIOR APPLICATION NUMBER: AU PRI381  
PRIOR FILING DATE: 2000-11-10  
PRIOR APPLICATION NUMBER: US 60/249,596  
PRIOR FILING DATE: 2000-11-17  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 622  
TYPE: DNA  
ORGANISM: Lawsonia intracellularis  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1) ... (621)  
US-10-010-160-1

Query Match 100.0%; Score 622; DB 14; Length 622;  
Best Local Similarity 100.0%; Pred. No. 2.3e-123;  
Matches 622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGATGATCCAGTAATAACGAGAAAGCAACCCGGAAGCAGCTGAGAAAGCTCGT 60  
DB 1 ATGCTGATGATCCAGTAATAACGAGAAAGCAACCCGGAAGCAGCTGAGAAAGCTCGT 60  
QY 61 TCTGAAGGAGGTGCTCCTTAATGAGAGAGCTTACTTAAGCATTTGACTGACAGAGG 120

Db	61	TCGTAAAGGGAAGTGCCTCCCTAAATCAGAAAGGTTACTAAACATTCATCTGCACAGAGG	120
Qy	121	ATGCTGGGGGCTTCGATTTATTCACAGCCGTAATGGGACGTCATTTTGAAACAATTTTCTAC	180
Db	121	ATGCTGGGGGCTTCGATTTATTCACAGCCGTAATGGGACGTCATTTTGAAACAATTTTCTAC	180
Qy	181	TATATTTTTCACGAATTCATTTTCGGTTTGAGTTACAGCACAGTCAGATATGCTTTATTT	240
Db	181	TATATTTTTCACGAATTCATTTTCGGTTTGAGTTACAGCACAGTCAGATATGCTTTATTT	240
Qy	241	ATTATGTTGCTCAAGAGATAGCTATTTATTTAGTACCAATTTTACTTTTATTTGCTGTT	300
Db	241	ATTATGTTGCTCAAGAGATAGCTATTTATTTAGTACCAATTTTACTTTTATTTGCTGTT	300
Qy	301	ACGGCATGGAATTCATTACGTGTACAGTTGGTGTGATTAATGACTACAAAGGTTTAA	360
Db	301	ACGGCATGGAATTCATTACGTGTACAGTTGGTGTGATTAATGACTACAAAGGTTTAA	360
Qy	361	TTTAAATGAGATTAATTAATTAATTAATAAAGGTTGAAGGAATGTTTCTTCTACAA	420
Db	361	TTTAAATGAGATTAATTAATTAATTAATAAAGGTTGAAGGAATGTTTCTTCTACAA	420
Qy	421	ACACTTGTTCGACTTTTACGTAGTTAGTTCACAGTAATGTATAGTATTTGCTCATAT	480
Db	421	ACACTTGTTCGACTTTTACGTAGTTAGTTCACAGTAATGTATAGTATTTGCTCATAT	480
Qy	481	ATGATTAATAAAGGAGGAGTTTCCAAACTTTTACCATTAATATTAATGCAAGTCTTCAGGT	540
Db	481	ATGATTAATAAAGGAGGAGTTTCCAAACTTTTACCATTAATATTAATGCAAGTCTTCAGGT	540
Qy	541	GTGGCAGATTATGCTTAATAACAGGAATAGACTGTTTATATATACGCTAATTTCCATG	600
Db	541	GTGGCAGATTATGCTTAATAACAGGAATAGACTGTTTATATATACGCTAATTTCCATG	600
Qy	601	ACAATTAATTCAGTGCAGATC 622	
Db	601	ACAATTAATTCAGTGCAGATC 622	

```

RESULT 2
US-10-311-455-118
Sequence 118, Application US/10311455
Publication No. US20030143606a1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis Associated with the Immune System by Determining Cytosine Methylation
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
CURRENT FILING DATE: 2002-12-16
PRIORITY APPLICATION NUMBER: PCT/EP01/07537
PRIORITY FILING DATE: 2001-07-02
PRIORITY APPLICATION NUMBER: DE 10032529.7
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: DE 10043826.1
PRIORITY FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 118
LENGTH: 9052
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-118

```

```
Query Match      8.8%; Score 54.8; DB 12; length 9052;  
Best Local Similarity   47.28%; Pred. No. 0.083;  
Matches 167; Conservative 0; Mismatches 187; Indels 0  
  
138 TTATTCAGGGCGGTAAATGGAACGCATTTTCTACTCATATTTTTACAGAATC 197
```

[illegible]

```

RESULT 3
US-09-895-913A-45/C
: Sequence 45: Application US/09895913A
: Patent No. US20020160456A1
: GENERAL INFORMATION:
: APPLICANT: Kleantous, Harold
: APPLICANT: Al-Garawi, Amal
: APPLICANT: Miller, Charles
: APPLICANT: Tomb, Jean Francois
: APPLICANT: Oomen, Raymond P.
: TITLE OF INVENTION: Identification of Polynucleotides
: TITLE OF INVENTION: Encoding No. US20020160456A1 Helicobacter Polypeptides in the
: FILE OF INVENTION: Genome
: FILE REFERENCE: 06132/043002
: CURRENT APPLICATION NUMBER: US/09/895,913A
: CURRENT FILING DATE: 2001-06-29
: PRIOR APPLICATION NUMBER: US 08/881,227
: PRIOR FILING DATE: 1997-06-24
: NUMBER OF SEQ ID NOS: 368
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 45
: LENGTH: 945
: TYPE: DNA
: ORGANISM: Helicobacter pylori
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (61)...(636)
: US-09-895-913A-45

```

Query March	8.5% ;	Score 53 ;	DB 10 ;	Length 945 ;
Best Local Similarity	46.9% ;	Pred. No. 0.084 ;		
Matches 201 ;	Conservative 0 ;	Mismatches 225 ;	Indels 3 ;	Gaps 1
QY	19	AAAACAGAGAAAGCAACCCCGAAGCAGCGTCAGAGCTCGTTCTGAAAGGAGTGTCCCT	78	
Db	845	AAACCCGAMCTCCCTAGCGCGGAAAAAATATCCAAAAACAGAGAGAAAGCAATGTGCTT	786	
QY	79	AAATCAGAAAGGTTACTTAAGCACTTACTACTCAGAGAGGAGTGGGGCTTGCTAT	138	
Db	785	AAGAGCATGGAAGTGTGTGGGGGTTTTGGGGTTATTGGCCGGGCTAATTAGATTTTTTGT	726	
QY	139	TATTCAGGCGTAAATGGGACGTCAATTTTGAACAATTTTCTACTATATTTTACAGATCA	198	
Db	725	TTTTTTTATATGATGGGTGGATGCTTTAGCGAAATGATGCGCAGTGTGTAAGA---T	669	
QY	199	TTTTCGGTTTAGGTTACAGACAGTACGATATGCTTATTTATTTATTTAGTGCACAGAG	258	
Db	668	TTTTCCCTAGATTTTCAGTAAGAAGGTTCCAGAGCTGTTTAAACCACTGGCTAAAGAC	609	
QY	259	ATAGCTAATTTTATTTGATGCAATATTTACTTTTATATGCTGTTACGCGATGGATTTTCATTA	318	

Db 608 ACTTTTATGCTTTTACCATTATTTATCATTTTAGTGGGCGCTTTTATCTAAT 549  
QY 319 CGTGACAGATTGGTGCATATGACATCAAGGTTTATTAATTTAAATGAGTAATTT 378  
Db 548 GCTTGCAATTTGGCTGGCTCTTTGGCCCTTAAGTCATGAGCTTAATTTCTTAATC 489  
QY 379 AATAATATAAAGGCTGAAGAAAGAAAGTTTCTTCAACAACAAGCTTGTGACTTTTA 438  
Db 488 AACCTTATCAATGCGTCAAAACCTTTTCTTAAAAAAGCTCTTGTGAGGAGTTTG 429  
QY 439 CGTAGTTTA 447  
Db 428 ATCACCCTTA 420

RESULT 4  
US-10-311-455-374  
; Sequence 374, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455  
; PRIOR FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 374  
; LENGTH: 7167  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-374

Query Match 8.5%; Score 52.8; DB 12; Length 7167;  
Best Local Similarity 46.3%; Pred. No. 0.2;  
Matches 174; Conservative 0; Mismatches 202; Indels 0; Gaps 0;

QY 234 TTATTTATTTATGCTGCTCAAGATAGCTATTTTATGATGCCAATTTACTTTTAT 293  
Db 709 TATATTTGTTAGTATTTAGAGATTTTAAAGTATAGATTTATTAATTTTATTTAG 768  
QY 294 TGCCTTACGCGATGATTTTATTAAGTATGATGATGATGATGATGATGATGATG 353  
Db 769 AGATTTAGTGTATGATTTTATTAATTTTATTAATTTTATTAATTTTATTAATTT 828  
QY 354 TTTTAAATTTTAAATGGAATTTTAAATTAATTAATTAATTAATTAATTAATTTGCTC 413  
Db 829 ATTATTTGTTAGATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 888  
QY 414 TCAACAACACTTGTGACTTTTACGATTTTACGATTTTACGATTTTACGATTTTACG 473  
Db 889 TGTGTTATATATGTTTAAATGTTTAGGGTTTTCGATGATTTGTTAGGTTTATTA 948  
QY 474 TCCATATATGTTTAAAGAGAGTTTCAACTTTTACCATTAATTTATGCAAGTCC 533  
Db 949 AATAATTTGCGGTTTAAAGCAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1008  
QY 534 TTCAGTGTGGCAGTATATGCTTAATATACAGAAATGATGATGATGATGATGATGAT 593  
Db 1009 TTTTGAAGTTTAAAGAAATTTATATTTTATGATTAATTTATGTTTTCGATTAATTTATG 1068  
QY 594 TCCATATGCAATTAAT 609

Db 1069 TTTTAAAGAAATTTT 1084

RESULT 5  
US-10-311-455-1532  
; Sequence 1532, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455  
; PRIOR FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 1532  
; LENGTH: 7128  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-1532

Query Match 8.4%; Score 52; DB 12; Length 7128;  
Best Local Similarity 46.5%; Pred. No. 0.3;  
Matches 240; Conservative 0; Mismatches 270; Indels 6; Gaps 2;

QY 78 TAAATCAGAAAGAGTTACTTAAGCATGACTGACAGAGAGTGGGCTTGTCTAT 137  
Db 4809 TAAATTTATGATTTATTTAATTTATTTAATTTAATTTAATTTAATTTAATTTAAT 4868  
QY 138 TTATTCAGCGTAAATGGGAGCTCAATTTGAACAATTTTCTACTATATTTTACAGATC 197  
Db 4869 TTAGTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4928  
QY 198 ATTTCGGTTGAGGTTACAGACAGATGATGATGATGATGATGATGATGATGATGATG 257  
Db 4929 TTTAATATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4988  
QY 258 GATAGCTATTTTATGATGCCAATTTACTTTTATGCTGTTACGCGATGATTTGAT 317  
Db 4989 TAT----TATGATGAAATATTTGAAGTTTATTTTATGAGTTAATTTAGTATTT 5043  
QY 318 ACGTGATCAAGTTGCTGATTTATGATGATGATGATGATGATGATGATGATGATG 377  
Db 5044 TAGTTTATTTATTTAGTATTTTATTTATTTATTTAATAA--TTATTTATGAGTATTTATTTATG 5102  
QY 378 TAAATTAATAAAGGTTGAAGAAATGTTTCTTCCACAACAACCTTGTGCACTTT 437  
Db 5103 TTGATATTTGTTTATTTATTTAGGATTAATAAGTAAATTAATTAATTAATTTATTTAT 5162  
QY 438 ACGTAGTTTATGTTCAAGTATTTATTTATGATTTTATTTATTTATTTATTTATTTAT 497  
Db 5163 GGGGTTTATTTTATTTAGTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 5222  
QY 498 GTTTCAACTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 557  
Db 5223 GATTTTATGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 5282  
QY 558 TAAATAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 593  
Db 5283 TAAATTTAATAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 5318

```

RESULT 7
US-10-311-455-2097
; Sequence 2097, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detecting
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455

```

Query Match	8.2%;	Score 51.2;	DB 12;	Length 12025;
Best Local Similarity	46.5%;	Pred. No. 0.54;		
Matches 201; Conservative	0;	Mismatches 228;	Indels 3;	Gaps 1;



QY 163 TTGAAACATTTCTACTATATTTTACAGAAATCTGGTTGAGGTACAGCAG 222  
DB TTTTAGAGATTTAGTTGATTTTATATTTATTCAGATGATGTTAGTTGAATTA 262  
QY 223 TCAGTATATGCTTATTTATTTATGCTCAAGAGATGCTATTTATTCAGTCCATA 282  
DB TAGAATATTTGGGTTTTTTTTTGTAAATGTTATGTTATTTTAAATATGATA 322  
QY 283 TTACTTTTATTTGCTGTACGAGATTTTCAATGATGATGAGTTGATGATTA 342  
DB TGTATTTTAGAGAGATTTTATGAGATTTGTTGATTAATTTATGATGTTGTTGA 382  
QY 343 ACTACAAAGTTTAAATTTAAATGAGTAAATTTATATATTAAGGGTTGAAAGA 402  
DB ATTAACGATATTTAAATGATGATGATTAATGATGATGATGATGATGATGATG 442  
QY 403 ATGTTGCTTCTCAACAAACTTG--TTGACTTTTACGATGTTAGTTCAAGTAA 459  
DB AAGTTTTTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 502  
QY 460 GTTATAGTATTTGCTCATATATATATTAAGAGAGTTTCAACTTTTACATTA 519  
DB TATTAATGATTTATTTAAATTTTATGATTAAGATTTTATGCTTTGATTAATGTA 562  
QY 520 TATATGCAAGTCTTCAGGTGTCAGATTAATGCTTAATACAGATGATGATGTT 579  
DB TTTTAAATTTAGTTGTTGTCGAATTAATGTTTATATATGATTAATTTAGTATG 622  
QY 580 TTATATACGCTA 591  
DB 623 TTAAAGTTGTA 634

## RESULT 9

US-10-311-455-1004  
Sequence 1004, Application US/10311455  
Publication No. US20030143606A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIEPENBROCK, Christian  
APPLICANT: BERLIN, Kurt  
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detecting  
FILE REFERENCE: 5013.1014  
CURRENT APPLICATION NUMBER: US/10/311.455  
PRIORITY FILING DATE: 2002-12-16  
PRIORITY APPLICATION NUMBER: PCT/EP01/07537  
PRIORITY FILING DATE: 2001-07-02  
PRIORITY APPLICATION NUMBER: DE 10032529.7  
PRIORITY FILING DATE: 2000-06-30  
PRIORITY APPLICATION NUMBER: DE 10043826.1  
PRIORITY FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 2424  
SEQ ID NO 1004  
LENGTH: 6118  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-1004

Query Match 8.2%; Score 51; DB 12; Length 6118;  
Best Local Similarity 50.5%; Pred. No. 0.46; Indels 3; Gaps 1;

Matches 151; Conservative 0; Mismatches 145; Indels 3; Gaps 1;

QY 227 TATATGCTTATTTATTTATGCTCAAGATGCTATTTTATGATGCAATATATAC 286  
DB 4654 TAAATAGCTATTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTT 4713  
QY 287 TTTTATTTGCTGATGATGATTTTATTTATTTATTTATTTATTTATTTATTTAT 346  
DB 4714 TTTTATTTGCTGATTTTATTTATTTATTTATTTATTTATTTATTTATTTAT 4770

QY 347 CAAGGTTTTTAAATTTAAATGAGTAAATTTATATATTAAGGGTTGAAAGAAAT 406  
DB 4771 GAGATTTTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 4830  
QY 407 TTGCTTCTCAACAAACTGTTGACTTTTACGATTTTACGATTTTACGATTTTAC 466  
DB 4831 AAGTTCATTTATTTGGGATTTAGTTTATTAATGATGATGATGATGATGAT 4890  
QY 467 GTATTTGCTATATATATATTAAGAGAGTTTCAACTTTTACATTAATATTA 525  
DB 4891 TTATGTTTATTTTATTTTGAATTAATGATTAATATATATATATATATATAT 4949

## RESULT 10

US-10-311-455-1944  
Sequence 1944, Application US/10311455  
Publication No. US20030143606A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIEPENBROCK, Christian  
APPLICANT: BERLIN, Kurt  
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detecting  
FILE REFERENCE: 5013.1014  
CURRENT APPLICATION NUMBER: US/10/311.455  
PRIORITY FILING DATE: 2002-12-16  
PRIORITY APPLICATION NUMBER: PCT/EP01/07537  
PRIORITY FILING DATE: 2001-07-02  
PRIORITY APPLICATION NUMBER: DE 10032529.7  
PRIORITY FILING DATE: 2000-06-30  
PRIORITY APPLICATION NUMBER: DE 10043826.1  
PRIORITY FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 2424  
SEQ ID NO 1944  
LENGTH: 7921  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-1944

Query Match 8.2%; Score 51; DB 12; Length 7921;  
Best Local Similarity 46.5%; Pred. No. 0.51; Indels 0; Gaps 0;  
Matches 165; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 159 TCATTTGAAACAATTTCTACTATATTTTACAGATTCATTTGCTTGAAGTTACAC 218  
DB 2153 TGAATTTTACAGCACTTATTTAGTTGTTAGTAAATTTTGTATTTTATTTTACGG 2212  
QY 219 ACAGTCAATATGCTTATTTATTTATGCTCAAGATGATGCTATTTATGATGCC 278  
DB 2213 TGTGGGGGATTTGATTTATTTATGATGTTTATTAATATTTATTTGATTAATG 2272  
QY 279 AATATATCTTTTATTTGCTGTTACGAGATTTCAATGATGATGATGATGATG 338  
DB 2273 AGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 2332  
QY 339 ATGACTACAAAGTTTAAATTTAAATGAGTAAATTTATTTATTAAGAGTTGAA 398  
DB 2333 TGGGAATTAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 2392  
QY 399 AGGAATGTTTGTCTTCAACAAACTGTTGCACTTTTACGATTTTATGATTAAT 458  
DB 2393 AAGATGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 2452  
QY 459 TGTATATGATTTGCTCATATATATTAAGAGAGTTTCAACTTTTATTA 513  
DB 2453 TATATATGTTGATTTTATTTATTTTATTTTGGGATGATTAATTTGTTATTTGTA 2507

## RESULT 11

US-10-311-455-1804  
Sequence 1804, Application US/10311455

Query Match	8.1%	Score 50.6	DB 12	Length 6126
Best Local Similarity	46.7%	Pred. No. 0.56		
Matches 161	Conservative	0	Mismatches 184	Indels 0
				Gaps 0

RESULT 12  
US-10-311-455-1958

```

? NUMBER OF SEQ ID NOS: 2424
? SEQ ID NO 1958
? LENGTH: 11047
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
? US-10-311-455-1958

```

Query Match	8.1%;	Score 50.6;	DB 12;	Length 11047;
Best Local Similarity	46.9%;	Pred. No. 0.71;		
Matches 158; Conservative	0;	Mismatches 179;	Indels 0;	Gaps 0

RESULT 13  
US-10-240

; Sequence 276, Application US/10240453  
; Publication No. US20030148326A1

```

; TYPE: DNA
; ORGANISM: Artificial Sequence

```

Query Match	Score	DB	Length
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens) US-10-240-453-276	8.1%	50.6	11047

Best Local Similarity 46.9%; Pred. No. 0.71;  
Matches 158; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY 164 TTGAACAACTTTTCTACTATATTTTACAGATCATCTTCGGTTAGGTTACAGACAGT 223  
DB 10382 TTGAATGAAATGAGTAATTTTGAATATATATTTGTTTGTAGTTAGTAGGA 10441  
QY 224 CAGTATGCTTATTTATTTATGTTGCTCAAGATAGTATTTATTTATGACCAATAT 283  
DB 10442 GCGATTTATGTTATTTAGTTATTTTGTGAATTTTGTAGTATTTTGTGGAAGTTGG 10501  
QY 284 TACTTTTATTTGCTGTTACGCGATGATTTCAATTCGTGTACCAAGTTGTCATTTATGSA 343  
DB 10502 TTTTATTTTAAAGTATGTAAGTTAAAGATTAAGATGATGATAGTGTGCTGTTTTTTT 10561  
QY 344 CTACAAAGCTTTTAAATTTAAATGAGTAAATTTATATATATTAAGGTTGAAGGAA 403  
DB 10562 TTGATTTATTTTATTTTATTTTAAATTTATTTTGTGCTTAAAGATTTGTAATATTA 10621  
QY 404 TGTTCCTTCTCAACAAACACTTGTTCGACTTTTACGTAGTTAGTTCAAGTATTTGTTA 463  
DB 10622 TTTTATTTATTTATTTATTTATTTTGTGTTAGTATGCGTATAGTTAGAGATTTGTTT 10681  
QY 464 TAGGTTATTTCCATATATATGATTTATTAAGAGAGAGTT 500  
DB 10682 TTGATGTAATTTTATTTAGATTTAGTTATTTGGGAATTT 10718

RESULT 14  
US-10-239-676-188  
; Sequence 188, Application US/10239676  
; Publication No. US20030082609A1

; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation  
; FILE REFERENCE: 5013.1003  
; CURRENT APPLICATION NUMBER: US/10/239,676  
; PRIOR FILING DATE: 2002-09-24  
; PRIOR APPLICATION NUMBER: PCT/EP01/03968  
; DE 10019058.8  
; DE 10019173.8  
; DE 10032529.7  
; DE 10043826.1  
; PRIOR FILING DATE: 2001-04-06  
; 2000-04-06  
; 2000-04-07  
; 2000-06-30  
; 2000-09-01  
; NUMBER OF SEQ ID NOS: 228  
; SEQ ID NO 188  
; LENGTH: 11047  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-239-676-188

Query Match 8.1%; Score 50.6; DB 14; Length 11047;  
Best Local Similarity 46.9%; Pred. No. 0.71;  
Matches 158; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY 164 TTGAACAACTTTTCTACTATATTTTACAGATCATCTTCGGTTAGGTTACAGACAGT 223  
DB 10382 TTGAATGAAATGAGTAATTTTGAATATATATTTGTTTGTAGTTAGTAGGA 10441  
QY 224 CAGTATGCTTATTTATTTATGTTGCTCAAGATAGTATTTATTTATGACCAATAT 283  
DB 10442 GCGATTTATGTTATTTAGTTATTTTGTGAATTTTGTAGTATTTTGTGGAAGTTGG 10501  
QY 284 TACTTTTATTTGCTGTTACGCGATGATTTCAATTCGTGTACCAAGTTGTCATTTATGSA 343  
DB 10502 TTTTATTTTAAAGTATGTAAGTTAAAGATTAAGATGATGATAGTGTGCTGTTTTTTT 10561  
QY 344 CTACAAAGCTTTTAAATTTAAATGAGTAAATTTATATATATTAAGGTTGAAGGAA 403  
DB 10562 TTGATTTATTTTATTTTATTTTAAATTTATTTTGTGCTTAAAGATTTGTAATATTA 10621  
QY 404 TGTTCCTTCTCAACAAACACTTGTTCGACTTTTACGTAGTTAGTTCAAGTATTTGTTA 463  
DB 10622 TTTTATTTATTTATTTATTTTGTGTTAGTATGCGTATAGTTAGAGATTTGTTT 10681  
QY 464 TAGGTTATTTCCATATATATGATTTATTAAGAGAGAGTT 500  
DB 10682 TTGATGTAATTTTATTTAGATTTAGTTATTTGGGAATTT 10718

DB 10502 TTTTATTTTAAAGTATGTAAGTTAAAGATTTAGTTAGTAAAGTGTGTTTTTTT 10561  
QY 344 CTACAAAGCTTTTAAATTTAAATGAGTAAATTTATATATTAAGGTTGAAGGAA 403  
DB 10562 TTGATTTATTTTATTTTATTTTAAATTTATTTTGTGCTTAAAGATTTGTAATATTA 10621  
QY 404 TGTTCCTTCTCAACAAACACTTGTTCGACTTTTACGTAGTTAGTTCAAGTATTTGTTA 463  
DB 10622 TTTTATTTATTTATTTATTTTGTGTTAGTATGCGTATAGTTAGAGATTTGTTT 10681  
QY 464 TAGGTTATTTCCATATATATGATTTATTAAGAGAGAGTT 500  
DB 10682 TTGATGTAATTTTATTTAGATTTAGTTATTTGGGAATTT 10718

RESULT 15  
US-09-754-853A-4/c  
; Sequence 4, Application US/09754853A  
; Publication No. US20030005491A1

; GENERAL INFORMATION:  
; APPLICANT: Hauge, Brian M.  
; APPLICANT: Parnell, Laurence D.  
; APPLICANT: Parsons, Jeremy D.  
; APPLICANT: Wang, Ming Li  
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With  
; FILE REFERENCE: 38-10(115810)B  
; CURRENT APPLICATION NUMBER: US/09/754,853A  
; PRIOR FILING DATE: 2001-01-05  
; PRIOR APPLICATION NUMBER: US 60/174,880  
; PRIOR FILING DATE: 2000-01-07  
; NUMBER OF SEQ ID NOS: 1119  
; SEQ ID NO 4  
; LENGTH: 513509  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (111805)..(113968), (114684)..(115204)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(513509)  
; OTHER INFORMATION: unsure at all n locations  
; OTHER INFORMATION: Clone ID: 318013\_region\_A3  
US-09-754-853A-4

Query Match 8.1%; Score 50.2; DB 11; Length 513509;  
Best Local Similarity 47.4%; Pred. No. 3.8;  
Matches 213; Conservative 0; Mismatches 233; Indels 3; Gaps 2;

QY 147 CGTAAATGGAGCTATTTTGAACAAATTTCTACTATATTTTAA-CAGATCATTTGCG 204  
DB 182633 CGTAAATGGAGAAATATTTTATACAAATTTGTATGACTTATTTTAAAGTAAATTTGATA 182574  
QY 205 TTGAGGTTACAGACAGTCAGTATATGCTTATTTATTTATTTGCTCAAGAGATGCT 264  
DB 182573 ATTTTACTGAAATAGATATATTTAAACTATTTTACTTAACATTTCAATTAACAAACC 182514  
QY 265 ATTTTATGATGCCAATATTAATTTTATTTGCTGTACCGCATGATTTCAATTCGTGTA 324  
DB 182513 TATTTATATAGTAATATGACTAATTTGTTAAACCTATATTTTGAACCGGTTGTTACGAATA 182454  
QY 325 CAAGTTGGTGATATATGACTACAAAGTTTAAATTTAAATGAGAGTAAATTTAATATA 384  
DB 182453 AAAAATGACATTTTATTTTACAGTATTTAGTATTTATTTAAAGATATCAAAATGAAAAA 182394  
QY 385 ATAAAGGTTGAAGAAATGTTTGTCTTCAACAAACACTTGTGACATTTTACGTAGT 444  
DB 182393 ATTAAGAAATTAATTAATGATTAATTAAGAGAAAGAAATGTTAGGAATATATTTTAT 182334  
QY 445 TTAGTTCAAGTATTTGTTAATGATTTGTTCCATATATATGATTTAAAGAGAGATTTTCA 504  
DB 182333 TTTTATTTATTTGATTAATTTCAATTAATTAATTAACCTTA-ATTTTATTTTACTTATGTTACTT 182275

QY 505 AACCTTTTACCATTAATTAATGCAAGTCTTCAAGTGTGCGAATTAATATGCTTAATAC 564  
DB 182274 AATTATGCGGTCAATTTTATTTTACCCTCATTCACAAATTAATAGCAATTAATACAA 182215  
QY 565 GGATAGTACTGTTTATTAATACGCTAAT 593  
DB 182214 ATTTTAGCAATATTTTCTCTATCTAAT 182186

## RESULT 16

US-10-239-676-133  
; Sequence 133, Application US/10239676  
; Publication No. US20030082609A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation  
; FILE REFERENCE: 5013.1003  
; CURRENT APPLICATION NUMBER: US/10/239,676  
; PRIOR FILING DATE: 2002-09-24  
; PRIOR APPLICATION NUMBER: PCT/EP01/03968  
; DE 10019058.8  
; DE 10019173.8  
; DE 10033529.7  
; DE 10043826.1  
; PRIOR FILING DATE: 2001-04-06  
; 2000-04-06  
; 2000-04-07  
; 2000-06-30  
; 2000-09-01  
; NUMBER OF SEQ ID NOS: 228  
; SEQ ID NO 133  
; LENGTH: 5935  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-239-676-133

Query Match 8.0%; Score 49.8; DB 14; Length 5935;  
Best Local Similarity 47.3%; Pred. No. 0.82;  
Matches 183; Conservative 0; Mismatches 202; Indels 2; Gaps 1;

QY 227 TATATGCTTATTTATTTATGTTGCTCAAGATACCTTTTATTGATGCCAATATTC 286  
DB 2608 TTTAAGATTTTATTTAGTGTGTTTGAGAGTGTGTTTATTTATTTATTAATTAAT 2667  
QY 287 TTTTATGCTGTGATGCGATGATTTCACTTACGTGTACAGTGTGATTAATGAGCTA 346  
DB 2668 TTTAATAGTAAATTTGTATGTTT--TTATATTAATGAATTTGTGAGAAATTAAT 2725  
QY 347 CAAAGCTTTTAAATTAATGAGTAATTAATTAATTAATTAAGGTTGAAGAATGT 406  
DB 2726 ATGAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2785  
QY 407 TTGCTTCGAACAAACCTGTTGCACTTTTACGTAAGTGAAGTAATTTGTAATG 466  
DB 2786 TAGATTTTAAATTAATTTTAAATTTTATTAATTAATTAATTAATTAATTTTAA 2845  
QY 467 GTATGTTCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 526  
DB 2846 ATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTAA 2905  
QY 527 CAAAGCTTCAGGTGTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 586  
DB 2906 GTTGAATTTGAATTTTATTTTGTGTTTATTAATTAATTAATTAATTAATTAAT 2965  
QY 587 CGCTAATTCCTATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 613  
DB 2966 TTGCAATTAATTAATTAATTTTGTG 2992

## RESULT 17

US-10-311-455-2045  
; Sequence 2045, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455  
; PRIOR FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 2045  
; LENGTH: 40862  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-2045

Query Match 8.0%; Score 49.8; DB 12; Length 40862;  
Best Local Similarity 45.2%; Pred. No. 1.7;  
Matches 183; Conservative 0; Mismatches 222; Indels 0; Gaps 0;

QY 108 TACTGACAGAGATGCTGGGCTTCTATTTATTCAGCGGTAAATGAGACGTCAATTTGA 167  
DB 16338 TATTGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 16397  
QY 168 AACATTTTCTACTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 227  
DB 16398 CGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 16457  
QY 228 ATATGCTTATTTATTTATTTATTTGCTCAAGATAGCTATTTATTAATTAATTAAT 287  
DB 16458 AATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 16517  
QY 288 TTTATGCTGTACGATGATTTCAATTAATTAATTAATTAATTAATTAATTAATTA 347  
DB 16518 TGGGTTTTTTTTTTTAAAGATTAATTAATTAATTAATTAATTAATTAATTA 16577  
QY 348 AAAGCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 407  
DB 16578 AATTTAGTGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 16637  
QY 408 TCTTCCTCAACAAACCTGTTGCACTTTTACGTAAGTGAAGTAATTTGTAATAG 467  
DB 16638 TATTAATTTTAAAGAAATTTTATTAATTAATTAATTAATTAATTAATTTTTC 16697  
QY 468 TATTTGCTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 512  
DB 16698 TAGTTTGTGTAATTAATTAATTTTGTGTTTATTAATTAATTAATTAATTTG 16742

## RESULT 18

US-10-311-455-1994  
; Sequence 1994, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ  
; FILE REFERENCE: 5013.1014



QY 342 GACTCAAGAGTTTAAATTAATGAGTAATTTATATATATAAAGGTTGAAG 401  
| | | | |  
Db 3293 GTTATTAAGGTTTATAGTAATAAGAGGTTGAATAAATGTTTATATATATGAAG 3352  
| | | | |  
QY 402 AATGTTTGCTTCAACAACACTTGTGACCTTTAGCTTATAGTCAAGTAATGT 461  
| | | | |  
Db 3353 AATGTTTATATGTTTATATATTTATTTATATAGATTTATATATATTTATATAT 3412  
| | | | |  
QY 462 TATAGTATTTGTCATATATATATATATATATATATATATATATATATATAT 510  
| | | | |  
Db 3413 TTTTATATTTGTTTATTTTATTTTATATGATTTGATTTTGAAGTT 3461  
| | | | |  
RESULT 21  
US-10-311-455-1893  
; Sequence 1893, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIERENBERG, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455  
; PRIOR FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 1893  
; LENGTH: 12590  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-1893  
Query Match 7.7%; Score 48.2; DB 12; Length 12590;  
Best Local Similarity 45.0%; Pred. No. 2.4; Indels 1; Gaps 1;  
Matches 220; Conservative 0; Mismatches 268;  
QY 80 AATCAGAGAGTTACTAAAGCATTTGACTACTGACAGAGGATGCTGGGCTTGTATTT 139  
| | | | |  
Db 752 ATTATAGTATTAATTAATTAAGAAATGATTTTGTATATAGTATGAAGTTCGGAATTTGA 811  
| | | | |  
QY 140 ATTGAGGCTAATGAGACGTCATTTTGAACAATTTCTACTATATTTTACAGATCAT 199  
| | | | |  
Db 812 AATGATGTTAATGGGTAATAAATTAAGAGAAATTTGTTTGTATTTTGAATTTTACG 871  
| | | | |  
QY 200 TTCGTTGAGGTACAGACAGTCAGTATATGCTTTAT-TTATTTATGCTGCTCAAG 258  
| | | | |  
Db 872 GATTTGTTGATTTTGTATTTGTTGTTATATATTTTATTTTATATGTTGCTTTG 931  
| | | | |  
QY 259 AATGCTATTATTAATGATGCAATATTAATTTATGCTGTTACGCGATGATTTCTATTA 318  
| | | | |  
Db 932 GAATTTTATTTTGTATTTATGTTTATATATTTTTCGTGCTGTTAAATTTTGTGTTG 991  
| | | | |  
QY 319 CGGTATCAAGTTGGTGCATTTATGACTACAAAGGTTTAAATTTAAATGAGTAAATTT 378  
| | | | |  
Db 992 TTTTATTTTGTATGAGAGATTAAGTTAATGTTTATTTTAAATTTTAAATTTAATA 1051  
| | | | |  
QY 379 AATATATATTAAGGTTGAAGGATGTTGCTTCAACAACACTTGTGCTGATTTTA 438  
| | | | |  
Db 1052 TTTGTAATAAATTTTATTTTAAATTAAGGATTTATTTATGATTTTACAGAGTATTTT 1111  
| | | | |  
QY 439 CGTATGTTTATGTTCAAGTAAATTTGTTATAGTATTTGCTCATATATGATTTTAAAGAGAG 498  
| | | | |  
Db 1112 GGCGGTTTGTATATTAATGATTTGTTAAATCGTATATATTTTGTGTTTAAATGATTA 1171  
| | | | |

QY 439 TTTTAACTTTTACATTAATATATGACAGTCTTCAAGTGGGAGATTTATAGCTT 558  
| | | | |  
Db 1112 TTTATTAAGTTTGTGTTTATTTTATTTTGTGAATATTTGATATGTTTAAATGTTGAA 1231  
| | | | |  
QY 559 AATACAGGA 567  
| | | | |  
Db 1232 TAGGTAGGA 1240  
| | | | |  
RESULT 22  
US-10-311-455-950  
; Sequence 950, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIERENBERG, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455  
; PRIOR FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 950  
; LENGTH: 18512  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-950  
Query Match 7.7%; Score 48.2; DB 12; Length 18512;  
Best Local Similarity 49.5%; Pred. No. 2.8; Indels 2; Gaps 1;  
Matches 152; Conservative 0; Mismatches 153;  
QY 168 AACATTTTCTACTATATTTTACAGAAATCAATTCGTTGAGTTACAGACAGTCACT 227  
| | | | |  
Db 4706 AATTTATAGTATTTTATGTTGATTTGATTAATTTGTTTGAAGTATATATAATTTTGT 4765  
| | | | |  
QY 228 ATATCTTATTTATTTATATGTTGCTCAAGATAGCTATTTTATGATGCAATATTTACT 287  
| | | | |  
Db 4766 TTTTATATGTTTATTTATTTTATTTTATGTTGTTATGTTATATATATATATATAT 4825  
| | | | |  
QY 288 TTTTATGCTGTTAGGAGATGATTTCAATTCGTTACAGATTTGTCATTTAGACTAC 347  
| | | | |  
Db 4826 GGTATTTGTTAATTAAGAAATGATTTTATATAGTTGTTTATATTTATG--TGG 4883  
| | | | |  
QY 348 AAAGTTTAAATTTAAATGAGATTAATTTAATTAATAAAGGTTGAAGAGATGTT 407  
| | | | |  
Db 4884 GAAAGGAAAGAGATTAATAAATTAATAATTAATATATGTTTATATATTTATTTA 4943  
| | | | |  
QY 408 TCGTTTCAACAACACTTGTGACCTTTTACGTATGTTTATGTTCAAGTATTTATAG 467  
| | | | |  
Db 4944 TGTATGTTATTTTATGAGGTTTATATTTTATATATTTTAAATTTGTTAANG 5003  
| | | | |  
QY 468 TATTTGT 474  
| | | | |  
Db 5004 TTTTITT 5010  
| | | | |  
RESULT 23  
US-10-311-455-1900  
; Sequence 1900, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander

APPLICANT: PIEPENROCK, Christian  
APPLICANT: BERLIN, Kurt  
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation  
FILE REFERENCE: 5013.1014  
CURRENT APPLICATION NUMBER: US/10/311.455  
PRIOR FILING DATE: 2002-12-16  
PRIOR APPLICATION NUMBER: PCT/EP01/07537  
PRIOR FILING DATE: 2001-07-02  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 2424  
SEQ ID NO 1900  
LENGTH: 13427  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-1900

Query Match 7.7%; Score 48; DB 12; Length 13427;  
Best Local Similarity 47.4%; Pred. No. 2.7;  
Matches 181; Conservative 0; Mismatches 195; Indels 6; Gaps 1;

QY 234 TTTATTTATTTATGTTGCTCAAGAGATAGCTATTTTATGATGCAATATTACTTTTAT 293  
DB 10874 TTTATTTATTTATTTGTTGTTAGGTAATATATTTGTAATATTTATTTATTTAT 10933  
QY 294 TGTGTTACGGCATGATTTGATTAAGTAAAGTTGTCATTATGAGTACAAAGT 353  
DB 10934 ATTAATTTTATGATTTGTTTAAATTTAGTATTTTAAATTAATAGTATTTA 10993  
QY 354 TTTT-----AAATTAATGAGTAATTTATTAATTAATAAGGTTGAAAGATCTT 407  
DB 10994 TTTTATGATTAATTTGAAATAGTAAATTAATTAATTAATTAATTAATTTT 11053  
QY 408 TGCTTCACAAACACTGTTGCTTTCAGTTCATGATTTAGTTCAGATTTGTTATAG 467  
DB 11054 TTTTATTAAGTAAGAAATTTGTTTATGATGTTGTAATGATTAATTAAGAGGTG 11113  
QY 468 TATGTTCCATATATGATTAATAAAGAGAGTTTCAACTTTTACCATATATATGTC 527  
DB 11114 AGTTTGTGTTTATTAAGATATATATATTTTGTAGATTTGATTTATGTTAAA 11173  
QY 528 AAGTCCTTACGTTGTCAGATTAATGCTTAATACAGAAATGACTGTTTATATAC 587  
DB 11174 AAGTTTATGAGTAATTAATTTATTCGTTGGAAATATATATTTAAATTAATGT 11233  
QY 588 GCTAATTCCTATGACATTTAT 609  
DB 11234 GGTGTTGATTTAGTATATATT 11255

RESULT 24  
US-10-311-455-80  
Sequence 80, Application US/10311455  
Publication No. US20030143606A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIEPENROCK, Christian  
APPLICANT: BERLIN, Kurt  
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation  
FILE REFERENCE: 5013.1014  
CURRENT APPLICATION NUMBER: US/10/311.455  
PRIOR FILING DATE: 2002-12-16  
PRIOR APPLICATION NUMBER: PCT/EP01/07537  
PRIOR FILING DATE: 2001-07-02  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 2424  
SEQ ID NO 80  
LENGTH: 19972  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-80

Query Match 7.7%; Score 47.8; DB 12; Length 19972;  
Best Local Similarity 46.4%; Pred. No. 3.5;  
Matches 199; Conservative 0; Mismatches 222; Indels 8; Gaps 1;

QY 181 TATATTTTACAGATGATTTGCGTTGAGGTTACACAGCTAGATATATCTTATTT 240  
DB 691 TTTTATTTTATGAGAGAAAGTAAATTTGTAATTTATTTATTTATTTATTTT 750  
QY 241 ATTTATGTTGCTCAAGAGATAGCTATTTTATGATGCAATATTACTTTATGCTGTT 300  
DB 751 ATATTAATTTGTTGTAAGATTAATTAATTTATTTATTTATTTATTTATTTT 810  
QY 301 AGGCATGATTTGATTAAGTAAAGTTGTCATTATGAGTACAAAGCTTTTAA 360  
DB 811 GAGATAGGCTTTTATTTATTTGTTAGGTTGAGTGAAGTGAATTTATTTT 868  
QY 361 TTTAATAGGAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 420  
DB 869 -----TGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 922  
QY 421 ACATGTTGCTTCAAGATTTGATTTTATGATTAATTTATTTATTTATTTT 480  
DB 923 GTTGATTTATGATTAAGTAAATTTATTTATTTATTTATTTATTTTATTT 982  
QY 481 ATGATTAATAAGAGAGTTTCAACTTTTATTTATTTATTTATTTATTTATTT 540  
DB 983 AGGTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1042  
QY 541 GTGCGATTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTT 600  
DB 1043 TTTTGGATTTGATGTTGATTTATTTATTTATTTATTTATTTATTTATTT 1102  
QY 601 ACAATTTAT 609  
DB 1103 GTAAATATT 1111

RESULT 25  
US-10-311-455-1530  
Sequence 1530, Application US/10311455  
Publication No. US20030143606A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIEPENROCK, Christian  
APPLICANT: BERLIN, Kurt  
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation  
FILE REFERENCE: 5013.1014  
CURRENT APPLICATION NUMBER: US/10/311.455  
PRIOR FILING DATE: 2002-12-16  
PRIOR APPLICATION NUMBER: PCT/EP01/07537  
PRIOR FILING DATE: 2001-07-02  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 2424  
SEQ ID NO 1530  
LENGTH: 7110  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-311-455-1530

Query Match 7.7%; Score 47.6; DB 12; Length 7110;  
Best Local Similarity 46.3%; Pred. No. 2.6;  
Matches 237; Conservative 0; Mismatches 264; Indels 11; Gaps 2;

```
OY 80 AATCCAGAGAGGTTACTAATAGCATGACACGACAGAGGATGCGGGCTTGCAATTT 139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5853 AAGTACAGATTTGTATTAATAGGATATATTAAGTTTAAAGGATTAATGATGATG 5912
OY 140 ATTCAGCGCTTAATGGACGCTCATTTTGAACAATTTCTACTATATTTTACAGATCAT 199
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5913 ATTAGTACGATATGATTAATTAATGATTTTAAATAGTATGATTTTAAATTTT 5972
OY 200 TTCCGTTTACGTTACAGACAGTC-----AGTATAGCTTTATTTATATGTTGCT 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5973 ATGTGTTATATTAATTAATTAATGTCGAAAGAAATTTTATTTTATTTTGAATAGC 6032
OY 253 CAAGAGATAGCTATTTTATTTGATGCGCATATTTCTTTTATTTGCTGTACGCGATGAT 312
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6033 GAATATGATTAATGATTTGATGATGATATTAATTTTATGATGTTTGTAGGTTAT 6092
OY 313 TCATTAAGTGTACAGTGTGTCATTTATGACTACAAAGGTTTAAATTTAAATGAGCT 372
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6093 TTGTTAAGATGTTTGTGTAATATTTTGAATTTTATAGGTTTATTTTATTTTAT 6152
OY 373 AATTTAATATATTAATTAAGGCTTGAAGAGATGTTGCTTCAACAACACTGTTGCA 432
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6153 AGGTGATTTTAAATTAAGGATTTAATTAATAGTTTAAATTTTAAATTTTGT 6212
OY 433 CTTTACGATGTTAGTTCAGATTAATTTGATGATTTTCCATATATGATTAATAA 492
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6213 ATTTTGTGTTTATTTATG-----TTTTTAATTAATATAGATTAATAATATATAA 6268
OY 493 GGAGAGTTTCAACTTTTACATTATATATGCAAGTCTTCAGGTGCGCATTTAT 552
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6269 ATTATTTGATTAATTAATTAATTTATTTTAAAGGTTTAAAGGTTTATTTTATTTTAT 6328
OY 553 ATGCTTAATACAGAACTACTACTGTTTATA 584
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6329 TTATTTTAAATTAATTAATTTATTTGTTTATTA 6360
```

RESULT 26

US-10-311-455-1486  
; Sequence 1486, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 1486  
; LENGTH: 11092  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-1486

Query Match 7.7%; Score 47.6; DB 12; Length 11092;  
Best Local Similarity 46.1%; Pred. No. 3.1;

Matches 201; Conservative 0; Mismatches 229; Indels 6; Gaps 1;

```
OY 162 TTTGAAACAATTTCTACTATATTTTACAGAAATCATTTGGTTGAGTTACAGACA 221
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10390 TTTTGTGAGGTTTGTGTTTGGGTTTGTATTTATTTGTTTGTGGAATATAT 10449
OY 222 GTCAGATTAATGCTTATATTTATTTATGTCGCAAGATAGCTATTTATGATGCCAT 281
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10450 GTTTTATTAAGATTTTATTTATTTTATTTAAATTAACGTTTATTAAGATATTA 10509
OY 282 ATTACTTTTATTTGCTGTACGCGCATGATTTCTTACGTTGATGATGATTTATG 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10510 TTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10569
OY 342 GACTCAAAAGGTTTAAATTTAAATGAGATTAATTAATTAATTAATTAATTAATTA 401
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10570 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10629
OY 402 AATGTTGCTTCTCAAC-----AACAAGTTGCGACTTTTACGATTTAGTCAAGT 455
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10630 ATATTATAGATTAATTAATTAATTAATTAATTTTGTTTTAAATTTTATTTAT 10689
OY 456 AATGTTATAGATTTGTTCAATATATATTAATTAATTAATTAATTAATTAATTA 515
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10690 TGTATTTTATTTATTTATTTAGTTATTTATTTTATTAATGATGATTAATTTAT 10749
OY 516 ATTTATTAATGCAAGTCTTCAGGTTGCGCATTAATGCTTAATACAGAAATGATCT 575
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10750 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10809
OY 576 TGTATTAATACGCTA 591
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10810 TATTTTGTGGTTA 10825
```

RESULT 27

US-10-311-455-999  
; Sequence 999, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 999  
; LENGTH: 6056  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-999

Query Match 7.6%; Score 47.4; DB 12; Length 6056;  
Best Local Similarity 45.7%; Pred. No. 2.7; Indels 0; Gaps 0;  
Matches 165; Conservative 0; Mismatches 196;

```
OY 161 ATTTGAACAATTTCTACTATATTTTACAGAAATCATTTGGTTGAGTTACAGAC 220
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3425 AATGCTAATTAATTAATTTGTATTTATAGTTGGAAGAAATTAATTTGTAATGAAAT 3484
OY 221 AGTCAGTATAGCTTTATTTATTTATGTTGCTCAAGAGATAGCTATTTTATGATGCCAA 280
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```





Db 3924 AATATTTGATTTTATTTTGGTTTTTTTTTTTAAATTAATAAATTTTGTAGG 3983  
QY 550 TATATGCTTAATACAGAAATAGTACTGTTTATTAATACGCTAAT 594  
Db 3984 TTGTTGTTTGTGTAAGTTTGTGTTTTTTTAAAGATAAT 4028

## RESULT 30

US-10-239-676-57  
; Sequence 57, Application US/10239676  
; Publication No. US20030082609A1  
; GENERAL INFORMATION:  
; APPLICANT: PIERENBERG, Christian  
; APPLICANT: BERLIN, Kurt  
; APPLICANT: PIERENBERG, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation  
; FILE REFERENCE: 5013.1003  
; CURRENT APPLICATION NUMBER: US/10/239,676  
; CURRENT FILING DATE: 2002-09-24  
; PRIOR APPLICATION NUMBER: PCT/EP01/03968  
; DE 10019058.8  
; DE 10019173.8  
; DE 10032529.7  
; DE 10043826.1  
; PRIOR FILING DATE: 2001-04-06  
; 2000-04-06  
; 2000-04-07  
; 2000-06-30  
; 2000-09-01  
; NUMBER OF SEQ ID NOS: 228  
; SEQ ID NO 57  
; LENGTH: 6385  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
; NAME/KEY: unsure  
; LOCATION: (5467)  
US-10-239-676-57

Query Match 7.6%; Score 47.4; DB 14; Length 6385;  
Best Local Similarity 46.7%; Pred. No. 2.8;  
Matches 217; Conservative 0; Mismatches 246; Indels 2; Gaps 2;

QY 131 TTGCTATTTATTCAGGCGTAATGAGAGCTCATTTGAAACAATTTCTAATATTTTA 190  
Db 3565 TTTTATTTTATTTGAATTTTATTTTATTTGATTTATTTTATTTTAAATTTAAGTTT 3624  
QY 191 CAGAAATCATTTGCGTTGAGTTAGACAGACAGTATAGCTTATTTATTTAATGTTG 250  
Db 3625 AATTTTATTTTAAATAGTTAAATTTTATTTTGAAGATAGATTTTATTTTATTTGTTG 3684  
QY 251 CTC-A-AGAGATAGCTATTTTATTTGATGCCAATATTTACTTTTATTTGCTGTACGGCATG 309  
Db 3685 TTTAATATGTTTTTATTTTGGTTTAAAGCATTTTGTGTTAAGTTTAAAGTG 3744  
QY 310 ATTTCATTCGTTGACAGTGTGTCATTATGACACAAAGTTTAAATTTAATG 369  
Db 3745 TTAGGATTAATGTTGTAATTTATTTAATTTTATTTTAAATGATTTTAAATGAAAAAT 3804  
QY 370 AGTAAATTTAATATAATAAAGGTTGAAGAGATGTTGCTTCCACAAACACTGTT 429  
Db 3805 TTTAATATTAATAAATAAAGAGAGTAATTTATTTTATTTTGAATTTAATATTTAT 3864  
QY 430 CGACTTTTACGTAGTTAGTCAAGTAATGTTAATAGTATGTCATATATGATTANA 489  
Db 3865 TTAATGATTA-GTTGTTTATTTAATATTTATTTTATTTTATTAATGTTAATTTAT 3923  
QY 490 AAGAGAGATTTTCAAACTTTTACATTAATATATGACAGTCTTCAAGGTGTGAGAT 549  
Db 3924 AATATTTGATTTTATTTTGTGTTTTTTTTTTTAAATTAATAAATTTTGTAGG 3983

QY 550 TATATGCTTAATACAGAAATAGTACTGTTTATTAATACGCTAAT 594  
Db 3984 TTGTTGTTTGTGTAAGTTTGTGTTTTTTTAAAGATAAT 4028

## RESULT 31

US-10-311-455-1822  
; Sequence 1822, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: PIERENBERG, Christian  
; APPLICANT: BERLIN, Kurt  
; APPLICANT: PIERENBERG, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determination of Cytosine Methylation  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 1822  
; LENGTH: 7057  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-1822

Query Match 7.6%; Score 47.4; DB 12; Length 7057;  
Best Local Similarity 50.4%; Pred. No. 2.9;  
Matches 170; Conservative 0; Mismatches 161; Indels 6; Gaps 2;

QY 167 AACAATTTTCTACATATTTTATACAGAAATCTTTGAGTTGAGTTACAGACAGTGC 226  
Db 3051 AATATATTTTATTTTGAATTTTGTGAATTTATTTATTTAAGATATGAGTTATATA 3110  
QY 227 TATATGCTTTATTTTATTTATGTTGCTCAAGATAGTATTTTATTTGATGCAATATTTAC 286  
Db 3111 TTAATGTTAATTTTATTTATGTTGCTAATAAATTAATGAATATAATATAATTAATA 3170  
QY 287 TTTTATTTGCTGTACGCGATGATTTCAATTCGTGTAACAAGTTGTCGATTTAGACTA 346  
Db 3171 TTTTAAATTTGATTAATGAATTAATTTGTAAGTGTGATGATTAATTTTGTG--- 3226  
QY 347 CAAGGTTTAAATTTAATGAGCTAATTAATATAATAAAGGTTGAAGAATGT 406  
Db 3227 -AAGGTTTGTATTAATTTTATTTGATTTTAAATTTAATGATGTTGGAATTT 3285  
QY 407 TTGCTTCAACAAACACTTG-TTCGACTTTTACGTAGTTTATTTCAAGTAATTTGTANA 465  
Db 3286 TTAATTTAATTTAATTTGTTTGAATTAATGATTAATAAATTTGTTGATGATTTTANA 3345  
QY 466 GGTATTTTCCATATATGATTAATAAAGAGAGTTT 502  
Db 3346 AATATTTTGAAGATTAATAAGAAAGTTAATAAT 3382

## RESULT 32

US-10-240-485-148  
; Sequence 148, Application US/10240485  
; Publication No. US20030148327A1  
; GENERAL INFORMATION:  
; APPLICANT: PIERENBERG, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Metastasis

FILE REFERENCE: 5013.1007  
CURRENT APPLICATION NUMBER: US/10/240.485  
CURRENT FILING DATE: 2002-10-02  
PRIOR APPLICATION NUMBER: PCT/EP01/03970  
PRIOR FILING DATE: 2001-04-06  
PRIOR APPLICATION NUMBER: DE 10019058.8  
PRIOR FILING DATE: 2000-04-06  
PRIOR APPLICATION NUMBER: DE 10019173.8  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 202  
SEQ ID NO 148  
LENGTH: 7057  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-240-485-148

Query Match  
Best Local Similarity 50.4%; Score 47.4; DB 12; Length 7057;  
Pred. No. 2.9; Mismatches 161; Indels 6; Gaps 2;

167 AACCAATTTCTACTATATTTTACAGATCATTTGGTTGAGTTACAGACAGTCAG 226  
3051 AAAAATTTTATTTTGAATTTTGGATTTTGAATTTTGAATTTTGAATTTTGA 3110  
227 TATATGCTTATTTATTTATTTATGCTTCAAGATAGCTATTTTATGATGCAATATTAC 286  
3111 TAAATGTTAATTTTATTTATGTTGTTAAATTAATTAATTAATTAATTAATA 3170  
287 TTTTATGCTGATGAGCAATGATTCATTCATGCTGATGATGCTGATGATGATG 346  
3171 TTTTAAATTTTGAATTAATTAATTTGATTAATTTGATTAATTTGATTAATTTG 3226  
347 CAAGATTTTAAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 406  
3227 -AAGGTTTGTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 3285  
407 TTGCTTCTCAACAAACCTTG-TTCGACTTTTACGATTTTATTTATTTATTTAT 465  
3286 TTAATTTATTTATTTAAATTTGTTTGAATTAATTAATTAATTTGTTTGAAT 3345  
466 GGAATTTGCTCATATGATTTATTAAGAGAGATTTT 502  
3346 AAATGTTTGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3382

## RESULT 33

US-10-311-455-1702  
Sequence 1702, Application US/10311455  
Publication No. US20030143606A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIEPENBROCK, Christian  
APPLICANT: BERLIN, Kurt  
TITLE OR INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect  
FILE REFERENCE: 5013.1014  
CURRENT APPLICATION NUMBER: US/10/311.455  
CURRENT FILING DATE: 2002-12-16  
PRIOR APPLICATION NUMBER: PCT/EP01/07537  
PRIOR FILING DATE: 2001-07-02  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 2424  
SEQ ID NO 1702  
LENGTH: 17721

TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-1702

Query Match  
Best Local Similarity 46.5%; Score 47.2; DB 12; Length 17721;  
Pred. No. 4.5; Mismatches 213; Indels 2; Gaps 1;

125 TGGGCTTGTCTATTTATTCAGGCTATGAGACCTCATTTGAAACAATTTCTACTATA 184  
11680 TGGTTTTTTTAAATTTATGATTAAGAAATTTTAAAGATTTTATGTTTATTA 11739  
185 TTTTACAGATCATTTGTTGAGTTACAGCA--CAGTCATATATGCTTATTTAT 242  
11740 TTTTATTAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 11799  
243 TTAATGCTCAAGATAGCTATTTTATTTATTTATTTATTTATTTATTTATTTAT 302  
11800 TATTTTATTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 11859  
303 GGCATGATTTTCATTTACGCTGACAGTTGCTGATTTATGACATTAAGCAAGTTTAAATT 362  
11860 AGAATAGTAAGAAATTAAGAAATTTTAAATTTTATTTATTTATTTATTTAT 11919  
363 TAAATGAGTAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 422  
11920 TATTTATGATGTTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 11979  
423 ACTTGTTCGCTTACGATTTTATTTATTTATTTATTTATTTATTTATTTATTT 482  
11980 GGTAAATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 12039  
483 GATTTATTAAGAGAGCTTTCAACTTTTATTTATTTATTTATTTATTTATTTAT 524  
12040 GAATATTAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 12081

## RESULT 34

US-10-311-455-1424  
Sequence 1424, Application US/10311455  
Publication No. US20030143606A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIEPENBROCK, Christian  
APPLICANT: BERLIN, Kurt  
TITLE OR INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect  
FILE REFERENCE: 5013.1014  
CURRENT APPLICATION NUMBER: US/10/311.455  
CURRENT FILING DATE: 2002-12-16  
PRIOR APPLICATION NUMBER: PCT/EP01/07537  
PRIOR FILING DATE: 2001-07-02  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 2424  
SEQ ID NO 1424  
LENGTH: 19787  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-1424

Query Match  
Best Local Similarity 50.0%; Score 47.2; DB 12; Length 19787;  
Pred. No. 4.7; Mismatches 143; Indels 12; Gaps 1;  
Matches 155; Conservative 0; Mismatches 143; Indels 12; Gaps 1;  
162 TTTTGAACAATTTCTACTATATTTTACAGATCATTTGGTTGAGTTACAGACACA 221



Db 3517 GGGTTAGGTATAGTGGTTTAAAGTTATTAATTTAGTAT 3554

RESULT 37  
US-10-240-453-269  
; Sequence 269, Application US/10240453  
; Publication No. US20030148326A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA  
; TITLE OF INVENTION: Transcription  
; TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated  
; FILE REFERENCE: 5013.1009  
; CURRENT APPLICATION NUMBER: US/10/240,453  
; CURRENT FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: PCT/EP01/03973  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: DE 10019058.8  
; PRIOR FILING DATE: 2000-04-06  
; PRIOR APPLICATION NUMBER: DE 10019173.8  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 350  
; SEQ ID NO 269  
; LENGTH: 15743  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-240-453-269

Query Match 7.5%; Score 46.8; DB 12; Length 15743;  
Best Local Similarity 46.5%; Pred. No. 5.2;  
Matches 185; Conservative 0; Mismatches 212; Indels 1; Gaps 1;

QY 129 GCTTGTATTTATTCAGCGGTAATGGAGCGTCAATTTGAAACAATTTTCACTAATATTT 188  
DB 12578 GTTTTATAGGTTTAAATGATTTTGTGTTAGTTTGTAGTGGGATTAATAGATA 12637

QY 189 TAGGAAATCAATTTGGGTTTGAGTTACGACAGTCGATATGCTTATTTATTTATGT 248  
DB 12638 TGTATTTATTAATTTAGTTAATTTTGTGTTATTTTGTAGTAGAATGGGTTTTTTATGT 12697

QY 249 TGCTCAAGAGATAGCTATTTTATTTGATGCAATATTTACTTTTATGCTGTACGGCATG 308  
DB 12698 TGGTATAGGTTGTTTAAATTTTCGATTTTGGATTTATTTGTTGGTTTTTTTAAAG 12757

QY 309 -GATTCATTACGTGTAACAAGTTGGTGCATTTAGACTACAAAGTTTTTAAATTTAAT 367  
DB 12758 TGTGAATATATATGATGATGATTTGGCTTTGGAGAGATGGGTTTTTATTTGAT 12817

QY 368 GGAATTAATTTAATATATAAAGGTTGAAGAAGATTTGCTTCTCAACAACACTTG 427  
DB 12818 TAGAAGTATTTGGTGTGATTAAGAGTGTGAGTGAAGGAATTTGCTTGTGTTT 12877

QY 428 TTCACATTTTACGATTTAGTTCAAGTAATTTGTTATGATTTGTCATATATGATTA 487  
DB 12878 TTTTATATGTTAATATATGATTTAAGGAAGTAATTTAGATTTGATGGGTTGTTATTT 12937

QY 488 TAAAGAGAGAGTTTCAACTTTTATCCATATATAT 525  
DB 12938 ATTATGATTTTTTAAATTTGTAAGATTAATATAT 12975

RESULT 38  
US-10-311-455-1593  
; Sequence 1593, Application US/10311455

Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ  
; TITLE OF INVENTION: Cytosine methylation  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 1593  
; LENGTH: 6254  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-1593

Query Match 7.5%; Score 46.6; DB 12; Length 6254;  
Best Local Similarity 46.0%; Pred. No. 4;  
Matches 157; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

QY 162 TTTTGAACAATTTTCTACTATATTTTATACAGATCAATTCGGTTGAGTTACAGACA 221  
DB 627 TTTTATGATATTTTAAATTTATTTTAAATATAGTGGAGTGATTTTAAAAATA 686

QY 222 GTGATATATGCTTATTTATTTATGTTGCTCAAGATATGCTATTTTATGATGCCAAT 281  
DB 687 TAAATTTGATTTTATGCTTATTTTGTGTTAAAGTTTGAATTTTTTTATGTTTTT 746

QY 282 ATTAATTTTATTTGCTGTACGCGCATGATTTTCAATGATGACAGTGGTGCATTATG 341  
DB 747 AGAATTTAATTTAATTTTATTTATTTATTTGATTTGATTTGATTTGTTAATATTA 806

QY 342 GACTACAAGGTTTTTAAATTTTAAATGAGTAATTTAATATTAATAAAGGTTGAAAG 401  
DB 807 GGATGAAGAAGTATGTTTATTAATATTTTAAATTTGTTTAAAGAAATGGAGAT 866

QY 402 AATGTTGCTCTCAACAACACTTTGCACTTTACGATTTAGTTAGTTCAAGTAATGT 461  
DB 867 TAGTTTTTATTTAGAAATTTTGTGAGAAAGGTGTTGTTATTTTAAATTTATGAAATAT 926

QY 462 TATAGTATTTGTCATATATGATTATTAAGAAGAGATTTT 502  
DB 927 ATTTGGTAATATATTTAAGATGTTTAAATTAATATATGTTT 967

RESULT 39  
US-10-204-708-11  
; Sequence 11, Application US/10204708  
; Publication No. US20030141852A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication  
; FILE REFERENCE: 5013.1012  
; CURRENT APPLICATION NUMBER: US/10/204,708  
; CURRENT FILING DATE: 2003-05-06  
; PRIOR APPLICATION NUMBER: PCT/EP01/03971  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: DE 10019058.8  
; PRIOR FILING DATE: 2000-04-06  
; PRIOR APPLICATION NUMBER: DE 10019173.8  
; PRIOR FILING DATE: 2000-04-07

PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 98  
SEQ ID NO 11  
LENGTH: 6317  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-204-708-11

Query Match 7.5%; Score 46.6; DB 12; Length 6317;  
Best Local Similarity 47.4%; Pred. No. 4.1;  
Matches 171; Conservative 0; Mismatches 189; Indels 1; Gaps 1;

QY 227 TATATGCTTATTTATTTATTTGCTCAAGAGATGCTATTTTATTTGATGCCAATATTAC 286  
DB 3143 TATACGATTAATTTGGGTTTAAATAATATTTTATTTTATTTATTTATTTTAA 3202  
QY 287 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 346  
DB 3203 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 3262  
QY 347 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 406  
DB 3263 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 3322  
QY 407 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 466  
DB 3323 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 3382  
QY 467 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 525  
DB 3383 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 3442  
QY 526 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 585  
DB 3443 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 3502  
QY 586 A 586  
DB 3503 A 3503

RESULT 40  
US-10-311-455-381  
Sequence 381, Application US/10311455  
Publication No. US20030143606A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: BERLIN, Kurt  
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detecting Cytosine Methylation  
FILE REFERENCE: 5013.1014  
CURRENT APPLICATION NUMBER: US/10/311.455  
PRIOR FILING DATE: 2002-12-16  
PRIOR APPLICATION NUMBER: PCT/EP01/07537  
PRIOR FILING DATE: 2001-07-02  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 2424  
SEQ ID NO 381  
LENGTH: 6317  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-381

Query Match 7.5%; Score 46.6; DB 12; Length 6317;  
Best Local Similarity 47.4%; Pred. No. 4.1;  
Matches 171; Conservative 0; Mismatches 189; Indels 1; Gaps 1;

QY 227 TATATGCTTATTTATTTATTTGCTCAAGAGATGCTATTTTATTTGATGCCAATATTAC 286  
DB 3143 TATACGATTAATTTGGGTTTAAATAATATTTTATTTTATTTATTTTAA 3202  
QY 287 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 346  
DB 3203 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 3262  
QY 347 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 406  
DB 3263 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 3322  
QY 407 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 466  
DB 3323 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 3382  
QY 467 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 525  
DB 3383 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 3442  
QY 526 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 585  
DB 3443 TTTTATTTGCTGTCAGCATGATTTCTAAGTGTGCAAGTTGGCATTATGACTA 3502  
QY 586 A 586  
DB 3503 A 3503

Search completed: October 4, 2003, 06:08:11  
Job time : 192 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 4, 2003, 02:48:32 ; Search time 1935 Seconds  
(without alignments)  
7812.602 Million cell updates/sec

```
Title: US-10-010-160-1
Perfect score: 622
Sequence: 1 atgtctgatgataccagtaa.....aattattgcagtcgcagatc 622
```

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Pre-processing:	Minimum Match	0%
Post-processing:	Minimum Match	0%

## Listing first 45 summaries

Database : EST:

21:	em_estbna.*
22:	em_esthna.*
23:	em_estin.*
24:	em_estmu.*
25:	em_estov.*
26:	em_estpl.*
27:	em_estro.*
28:	em_hrc.*
29:	gb_estl.*
30:	gb_est2.*
31:	gb_hrc.*
32:	gb_est3.*
33:	gb_est4.*
34:	gb_est5.*
35:	em_estfm.*
36:	em_estom.*
37:	em_gss_hum.*
38:	em_gss_inv.*
39:	em_gss_pln.*
40:	em_gss_vrt.*
41:	em_gss_fm.*
42:	em_gss_mam.*
43:	em_gss_mus.*
44:	em_gss_pro.*
45:	em_gss_rod.*
46:	em_gss_pbg.*
47:	em_gss_vrl.*
48:	gb_gss61.*
49:	gb_gss62.*

**Pred. No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description
No.							
C 1	60.4	9.7	996	29	CNS000FTH		AL071063 Drosophila
C 2	60.2	9.7	1101	29	CNS000KES		AL077453 Drosophila
C 3	59.2	9.5	1154	13	BX436282		BX436282 BX436282
C 4	59.2	9.5	1200	13	EX415878		EX415878 BX415878

C	5	58.6	9.4	886	28	BH1772277	BH1772277 008 L 228
C	6	58.6	9.4	886	29	CNS07JUX	AL614235 T3 end of
C	7	57.8	9.3	1201	13	BX360615	BX360615 BX360615
C	8	57.8	9.3	1201	13	BX424465	BX424465 BX424465
C	9	57.4	9.2	568	13	BU497346	BU497346 PFEST0ab6
C	10	57.4	9.2	1201	13	BX439779	BX439779 BX439779
C	11	56.8	9.1	563	13	BU495402	BU495402 PFEST0ab7
C	12	56.8	9.1	1101	29	CNS017V2	AL109536 Drosophill
C	13	56.6	9.1	734	29	CNS010MP	AL099163 Drosophill
C	14	56.6	9.1	1101	29	CNS000B8	AL063632 Drosophill
C	15	56.6	9.1	1201	13	BX461824	BX461824 BX461824
C	16	56.2	9.0	1201	13	AL565455	AL565455 AL565455
C	17	56	9.0	1146	29	AL565455	AL565455 AL565455
C	18	55.8	8.9	1101	29	CNS003DO	AL0645246 BX462546
C	19	55.4	8.9	977	13	BX415819	AL0645580 Drosophill
C	20	55.2	8.9	977	13	BQ452226	BX415819 BX415819
C	21	54.8	8.8	945	29	CNS04D0K	BQ452226 PFEST0a9
C	22	54.8	8.8	1201	13	AL536104	AL285149 Tetracodon
C	23	54.6	8.8	1201	13	CNS00EVL	AL536104 AL536104
C	24	54.4	8.7	994	13	BX414650	AL0659706 Drosophill
C	25	54.4	8.7	1092	29	CNS020K7	BX414650 BX414650
C	26	54.2	8.7	1021	29	CNS014DY	AL175696 Tetracodon
C	27	53.8	8.6	854	9	AL522840	AL104032 Drosophill
C	28	53.8	8.6	1076	13	BX338020	AL522840 AL522840
C	29	53.8	8.6	1201	29	CNS0167M	BX338020 BX338020
C	30	53.6	8.6	1101	29	CNS014XJ	AL106339 Drosophill
C	31	53.4	8.6	887	13	BX441520	AL104737 Drosophill
C	32	53.2	8.6	1201	13	BX41520	AL106339 Drosophill
C	33	52.8	8.5	757	29	BX214654	BX441520 BX441520
C	34	52.8	8.5	1056	13	BX415058	BX357585 BX357585
C	35	52.8	8.5	1101	29	CNS016L1	BX214654 Danilo ter
C	36	52.6	8.5	999	13	BX380865	BX415058 BX415058
C	37	52.6	8.5	1200	29	CNS016C0	AL106896 Drosophill
C	38	52.4	8.4	739	29	CNS06QXV	BX380865 BX380865
C	39	52.4	8.4	1101	29	CNS00DE7	AL106578 Drosophill
C	40	52	8.4	928	29	CNS00DKY	AL411257 T7 end of
C	41	52	8.4	1094	29	CNS012F2	AL069440 Drosophill
C	42	52	8.4	1177	13	BX399460	AL071865 Drosophill
C	43	52	8.4	1201	13	BX381083	AL101513 Drosophill
C	44	52	8.4	1201	13	BX426629	BX399460 BX399460
C	45	51.8	8.3	1045	13	BX456814	BX381083 BX381083
							BX426629 BX426629
							BX456814 BX456814

## ALIGNMENTS

RESULT 1  
 CINS00FUH/c  
 LOCUS  
 DEFINITION  
 CINS00FUH 996 bp DNA linear GSS 03-JUN-1999  
 Drosophila melanogaster genome survey sequence ITT3 end of BAC:  
 BACR1021 of RCI1-98 library from Drosophila melanogaster (fruit  
 fly), genomic survey sequence.  
 AL071063  
 AL071063.1 GI:4951105  
 GSS.  
 Drosophila melanogaster (fruit fly)  
 Drosophila melanogaster  
 Eukaryote; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Dipera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 996)  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 Direct Submission  
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
 BP 101 91006 EVRY cedex - FRANCE (E-mail : seqret@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)  
 Determination of this BAC-end sequence was carried out as part of a  
 collaboration with the Berkeley Drosophila Genome Project (BDGP).  
 The BDGP is constructing a physical map of the Drosophila  
 melanogaster genome using these BACs. For further information  
 please see <http://www.fruitfly.org> The BDGP Drosophila  
 melanogaster BAC library was prepared by Kazutoyo Osogawa and  
 Aaron Mammosser in Pletier de Jong's laboratory in the Department of

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

## FEATURES

source

Location/Qualifiers

1..996

/organism="Drosophila melanogaster"

/mol\_type="genomic DNA"

/db\_xref="taxon:7227"

/clone="BACR31021"

/clone\_11b="RPCI-98"

/note="end : TET3"

BASE COUNT 383 a 164 c 81 g 171 t 197 others

ORIGIN

Query Match 9.7%; Score 60.4; DB 29; Length 996;

Best Local Similarity 34.5%; Pred. No. 0.1; Mismatches 223; Indels 0; Gaps 0;

Matches 154; Conservative 69; Mismatches 223; Indels 0; Gaps 0;

164 TGAAGACATTTCTACATATTTTACAGATTCATTCGTTGAGTTACAGACAGT 223

962 TTTMAAAATTTMAATTTTATTTTATTTTBTBSSSSSSSSSSSSSSGRRAAA 903

224 CAGTATATGCTTATTTATTTATTTGCTCAAGATACCTATTTATGATCCAAAT 283

902 AAAAAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 843

284 TACTTTTATTTGCTTACGATGATTCATTTACGTTACAGTGTGATTTATGGA 343

842 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 783

344 CTACAAAGCTTTTAAATTTAAATGAGTAAATTTTAAATTTAAATGAGTAAAGGA 403

782 AATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 723

404 TGTTCCTTCTCAACAAACCTTTGACCTTTACGATTTAGTTAGTTACGATTTGTTA 463

722 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 663

464 TAGGATTTGCTCATATGATTTAAAGAGAGTTTCAACTTTTACCATTTATAT 523

662 AAACCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 603

524 ATGCAAGTCTTCAGGTGTCAGATTTATGCTTATATCAGGAATAGTACTTTTAT 583

602 AATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 543

584 ATACGCTAATTCCTATGACATTTAT 609

542 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 517

RESULT 2

CNS00KB5 1101 bp DNA 1linear GSS 03-JUN-1999

LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC:

DEFINITION BAC39P05 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL077453 GI:4956930

VERSION AL077453.1 GI:4956930

KEYWORDS GSS.

SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydriidae; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1101)

AUTHORS Genoscope.

TITLE Direct Submission

## JOURNAL

## COMMENT

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 91006 EVRY cedex - FRANCE (E-mail : [sequef@genoscope.cns.fr](mailto:sequef@genoscope.cns.fr) Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr))

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Osoegawa and Aaron Mosmoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

## FEATURES

source

Location/Qualifiers

1..1101

/organism="Drosophila melanogaster"

/mol\_type="genomic DNA"

/db\_xref="taxon:7227"

/clone="BACR39P05"

/clone\_11b="RPCI-98"

BASE COUNT 540 a 0 c 139 g 252 t 170 others

ORIGIN

Query Match 9.7%; Score 60.2; DB 29; Length 1101;

Best Local Similarity 42.3%; Pred. No. 0.11;

Matches 185; Conservative 10; Mismatches 242; Indels 0; Gaps 0;

173 TTTCTACATATTTTACAGATTCATTCGTTGAGCTTACAGACAGTCAGATATG 232

578 TTTTNTTTNTTTTNTTTTNTTTTATTTTNTTTTNTTTTNTTTTNTTTTNTTT 519

233 CTATTTATTTATGCTGCTCAAGATAGCTATTTTATGATGCAATATTTACTTTTA 292

518 TTTTAAAAAATTTTNTTTTAAATTAATTTNTTTTNTTTNTTTNTTTNTTTNT 459

233 TTGCTGTTCAGCATGATTCATTTACGTTGACAGATTTGATGACTACAAAGG 352

458 TTTTNTTTTATTTTATTTTNTTTTATTTTNTTTTATTTTNTTTTNTTTTATTT 399

333 TTTTAAATTTAAATGAGTAAATTTATTAATTAAGGTTGAAGAAATGTTGCT 412

398 TTTTAAAAATTTAAAAATTTAAAAATTTAAATTTNTTTATTAATTTNTTTT 339

413 CTCAACAAACCTTTGTCAGCTTTACGTTAGTTCAAGTATTTGTTATGATATG 472

338 ATTNNNNNNNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNT 279

473 TTCCATATATGATTTAAAGAGAGTTTCAACTTTTACCATTAATTTATGCAAG 532

278 TTTTNTTTTATTTTNTTTTATTTTNTTTTNTTTTNTTTTNTTTTNTTTTNT 219

533 CTTCAGGTGTCAGATTTATGCTTATATACAGAAATGTAAGTTTATATACCTTA 592

218 WTATATTTTAAATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 159

593 TTCCATATGCAATTTAT 609

158 TTTATTTATTTTNTTTT 142

RESULT 3

BX436282 1124 bp mRNA 1linear EST 15-MAY-2003

LOCUS BX436282 Homo sapiens THYMUS Homo sapiens cDNA CSOCAP001YC01

DEFINITION 5-PRIME, mRNA sequence.

ACCESSION BX436282

VERSION BX436282.1 GI:30787521



KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS	Li, W.B., Gruber, C., Jesssee, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0CAP001AB010P1. location/Qualifiers
FEATURES	1. 1124
source	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0CAP001YCO1" /tissue_type="THYMUS" /clone_lib="Homo sapiens THYMUS" /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
BASE COUNT	129 a 29 c 109 g 435 t 422 others
ORIGIN	
Query Match	9.5%; Score 59.2; DB 13; Length 1124;
Best Local Similarity	23.8%; Pred. No. 0.17;
Matches 102;	Conservative 139; Mismatches 188; Indels 0; Gaps 0;
144 AGCGTAATGGACGTCATTGTAACAATTTCTACTAATTTTACAGAAATCATTTG	203
:: ::	
694 RRRRRRRRRRRDTTTTTTTTTDDRDTTTTTTTTTDDTTTTTTTTRRRRDTTTTT	753
204 GTTGAGGTGACAGACAGTCAGTATGCTTATTATTTAATGTTGTCAGAGATAGC	263
:: ::	
754 TTTTDDTTRRRRRRTTTRRRRRRTTTRRRRRRTTTRRRRRRTTTRRRRRRRRR	813
264 TATTTATGATGCAATATTAATCTTTATGCTGTAACGATGATTCATTCAGTGT	323
:: ::	
814 RRTTTTTTTTTTTRRRRRRRRRRTTTTTTTTTTTTTTTTTTTTDDRRRRRTTTTT	873
324 ACAAGTGTGCATTATGACATCAACAAGTTTTTAATTAATGAGATAATTTAAT	383
:: ::	
874 TDDRRRRRRRRDTTTTTTTTTDDTTTTTDDDDTTTTTDRRRRTTRRRRRRRRR	933
384 AATAAAGGTTGAAGAAGATGTTGCTTCACAACAACACTGTTGACCTTTACGTA	443
:: ::	
934 RRRRRRTTDDRRRRRTTTTTTTTTTTRRRRRRDDDKRRRTTXXTTTTTTTTT	993
444 TTATGTAAGTAATTTATATGATATGTTCCATATATGATTAATAAAGAGAGATTTC	503
:: ::	
994 TTRRRRRRRRRRTTDDRRRRRRRRRRRDTTTTTTTTTTRRRRTTRRRRRRTTTT	1053
504 AAACCTTTTACATTAATTAATGCAAGTCCTTCAGGTGCGAGATTAATGCTTAAT	563
:: ::	
1054 DRRRRRTTTRTTRDTRDRTTTTTTTTTTTTTTTRRRRRRDDDTTTTTTTTTTTT	1113
564 AGCAATACT 572	
DB	1114 RRRRRRTT 1122
LOCUS	BX415878
RESULT 4	1200 bp
BX415878/c	mRNA
	linear
	EST 15-MAY-2003

DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
EX415878 Homo sapiens THYMUS	EX415878	EX415878	5-PRIME, mRNA sequence.	EST.	Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	1 (bases 1 to 1200)	I.J.W.B., Gruber, C., Jessee, J. and Potayes, D.	Unpublished	Genoscope -- Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: segret@genoscope.cns.fr; Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifestech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0CAP008B020P1.
										Location/Qualifiers 1..1200

FEATURES					
	SOURCE	Location/Qualifiers			
		1..1200			
		/organism="Homo sapiens"			
		/mol_type="mRNA"			
		/db_xref="taxon:9606"			
		/clone="CS0CAP008YI04"			
		/issue_type="THYMUS"			
		/clone_lib="Homo sapiens THYMUS"			
		/notes=Vector: pCMVSPORT6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector.			
		Library was not normalized."			
BASE COUNT	428 a	70 c	85 g	447 t	170 others
ORIGIN					
Query Match	9.5%; Score 59.2;	DB 13;	Length 1200;		
Best Local Similarity	37.2%;	Pred. No. 0.17;			
Matches 158;	Conservative 59;	Mismatches 207;	Indels 1;	Gaps 1;	
Oy	162	TTTGAAACAATTTCTACTATATTTTTACAGAAATCATTCGGTTGAGGTACAGACA	221	:	:
Db	1081	WTWAAAAMATTATTMTAAATATATAWMAAAAMTWTAAMAAAAATAATTTTTTTTAAWMM	1022	:	:
Oy	222	GTCAGTANAGCCTTATTATTATNTAGTGCACAGAATGCTATTTATATGAGCCAT	281	:	:
Db	1021	AAATATAMWTTTAAAMWMTAAWMAWTTTAAWMAWTTTAAWMTAANNATTTATATAMWTTWWMA	962	:	:
Oy	282	ATTACTTTTTTATGCTGTACCGCATGATTCATTACGTGACAGTTGGCATTTAG	341	:	:
Db	961	ATTTTTTTTWTWTATTTTATTTATTTAWMAAAAMATTTMAAAATTTMAAAATTAAMAA	902	:	:
Oy	342	GACTACAAAGCTTTTAAATTTAAATGAGTAAATTAATTAATTAATAAAGGCTGAAG	401	:	:
Db	901	AAMTTTTTSTATTATTMTTAAATTAATTAATTAATAAATTAATAAATTTTTTTTAAAT	842	:	:
Oy	402	AATGTTGGCTCCACAAACAACCTGTTCGACCTTTACGTAGTTAGTACAGTAATGT	461	:	:
Db	841	WTTTTTAAATTAAMAAMTTTTTAAAMTTTTTTTTTAAATTTTGTAAWTTWMAAARATTA	782	:	:
Oy	462	TATAGGTAT-TGTTCCATATATGATTTATAAAGAGAGATTTCAAACCTTTTACCATTAT	520	:	:
Db	781	TAMATWATATTTTATGHTWATAATATATATATATATATATATATATTTCTTTWMAAAAM	722	:	:
Oy	521	ATTATGCAAGCTTCACAGGTGCGACATTAATATGCTTAATACGAATAGTACTGTT	580	:	:
Db	721	WWWATTTAAATTAATAAATAATATATATATATTAATCTAATTAATAAATAAAMATATAMTTWT	662	:	:
Oy	581	TATAT 585	:	:	:
Db	661	YWAMT 657	:	:	:

```
RESULT 5
BH177277/c 886 bp DNA linear GSS 19-OCT-2001
LOCUS 008 L 22-rev SmbAC1 Schistosoma mansoni genomic clone 008L22 5',
DEFINITION BH177277
ACCESSION BH177277
VERSION BH177277.1 GI:16275873
KEYWORDS GSS.
SOURCE Schistosoma mansoni
Schistosoma mansoni
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigeida; Schistosomatidae; Schistosomatidae; Schistosoma.
1 (bases 1 to 886)
Le Paslier,M.-C., Pierce,R.J., Merlin,F., Hirai,H., Wu,W., Williams
,D., Johnston,D., Loverde,P.T. and Le Paslier,D.
Construction and characterization of a Schistosoma mansoni
bacterial artificial chromosome library
Genomics 65 (2), 87-94 (2000)
JOURNAL MEDLINE
20247247
10783255
COMMENT Other GSSs: 008 L 22-21
INSERM U 167
Contact: Pierre RJ
Institut Pasteur de Lille
1 rue du Professeur A. Calmette, 59019-Lille, France
Tel: (33) (0)3 2087783
Fax: (33) (0)3 2087788
Email: Raymond.Pierce@pasteur-lille.fr
CNS sequencing ID=DCGA008DP11BP1
Plate: 008 row: L column: 22
Seq primer: M13 reverse primer
Class: BAC ends
High quality sequence stop: 886.
Location/Qualifiers
1. 886
/organism="Schistosoma mansoni"
/mol_type="genomic DNA"
/strain="Puerto-Rican"
/db_xref="taxon:6183"
/clone="008DP11"
/sex="mixed"
/dev_stage="cercariae"
/lab_host="Biomphalaria glabrata"
/clone_lib="SmbAC1"
/Note="Vector: pBelOBAC 11; Site 1: Hind III; Partially
Hind III digested and size-selected S. mansoni cercarial
DNA was ligated into Hind III digested pBelOBAC 11 vector
and used to transform E. coli DH10B. The complete library
contains 23808 clones from 4 independent
sizing-ligation-transformations. Average insert size
ranges from 70-127 kb and genome coverage is 7.9-fold."
BASE COUNT 376 a 22 c 49 g 288 t 151 others
ORIGIN
Query Match 9.4%; Score 58.6; DB 28; Length 886;
Best Local Similarity 40.6%; Pred. No. 0.22;
Matches 176; Conservative 50; Mismatches 199; Indels 9; Gaps 2;
QY 173 TTTTACATATATTTTACAGATCATTTGCGTTGAGTTACAGACAGTCAGTATATG 232
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 833 TTTTATTTTATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTATTT 774
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 233 CTTTATTTATTTATGTTGCTCAAGATAGCTATTTATTTATGATGCAATATTACTTTT 292
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 773 TTTTATTTTATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 714
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 293 TTGCTGTACGAGATGATTTCTTACGTACAGTGTGTCATTATGACCTACAAAG 352
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 713 TATTTTATTTATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 661
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 353 TTTTAAATTTAATGAGTAATTTTATTTAATTTAATTTAATTTAATTTAATTTGCTT 412
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
```

```
DB 660 TTTTAAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTTWA 601
QY 413 CTCAACAAACACTGTTGCACTTTAGCTAGTTAGTTCAAGTAATGTTATGATTTG 472
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 600 TWAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTTWA 541
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 473 TTCCATATATGATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTG 532
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 540 TTTTAAATTTAATTTTAKAAAAAAATTTTAKTTTAKTTTAKTTTAKTTTAKTTT 483
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 533 CTTCAGGTGTCAGATTTATGCTTTATATACAGATTTATGATTTATTTATTTATG 592
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 482 TKTAAATTTATTTTAAATTTAATTTAATTTAATTTAATTTAATTTAATTTATTT 423
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 593 TTTCCATGACATTT 606
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 422 TKTATTTTAKTTT 409
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 6
CNS07JUX/c 886 bp DNA linear GSS 30-NOV-2001
LOCUS T3 end of clone 008DP11 of library SmbAC1 from strain Puerto-Rican
DEFINITION T3 end of clone 008DP11 of library SmbAC1 from strain Puerto-Rican
ACCESSION AL614235
VERSION AL614235
KEYWORDS GSS.
SOURCE Schistosoma mansoni
Schistosoma mansoni
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigeida; Schistosomatidae; Schistosomatidae; Schistosoma.
1 (bases 1 to 886)
Le Paslier,M.-C., Pierce,R.J., Merlin,F., Hirai,H., Wu,W.,
Williams,D.L., Johnston,D., Loverde,P.T. and Le Paslier,D.
Construction and characterization of a Schistosoma mansoni
bacterial artificial chromosome library
Genomics 65 (2), 87-94 (2000)
JOURNAL MEDLINE
20247247
10783255
COMMENT Direct Submission
Submitted (05-OCT-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequefgenoscope.cns.fr)
Web : www.genoscope.cns.fr)
Partially Hind III digested and size-selected S. mansoni cercarial
DNA was ligated into Hind III digested pBelOBAC 11 vector and used
to transform E. coli DH10B. The complete library contains 23808
clones from 4 independent sizing-ligation-transformations. Average
insert size ranges from 70-127 kb and genome coverage is 7.9-fold.
FEATURES
1. 886
Location/Qualifiers
/organism="Schistosoma mansoni"
/mol_type="genomic DNA"
/strain="Puerto-Rican"
/db_xref="taxon:6183"
/clone="008DP11"
/clone_lib="SmbAC1"
/Note="end : T3"
BASE COUNT 376 a 22 c 49 g 288 t 151 others
ORIGIN
Query Match 9.4%; Score 58.6; DB 29; Length 886;
Best Local Similarity 40.6%; Pred. No. 0.22;
Matches 176; Conservative 50; Mismatches 199; Indels 9; Gaps 2;
QY 173 TTTTACATATATTTTACAGATCATTTGCGTTGAGTTACAGACAGTCAGTATATG 232
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 833 TTTTATTTTATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTATTT 774
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 233 CTTTATTTATTTATGTTGCTCAAGATAGCTATTTTATTTATGATGCAATATTACTTTT 292
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 773 TTTTATTTTATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 714
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
```

```

FEATURES
    source
        location/Qualifiers
            1..1201
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0DE002YP24"
                /tissue_type="PLACENTA"
                /clone_lib="Homo sapiens PLACENTA"
                /note="Vector: pCMVSPORT 6: 1st strand cDNA was primed
                    with a NotI-oligo(dT) primer. Five prime end enriched,
                    double-strand cDNA was digested with Not I and cloned into
                    the Not I and EcoRV sites of the pCMVSPORT 6 vector."

```

BASE COUNT 508 a 103 c 175 g 319 t 96 others  
 ORIGIN

Query Match 9.3%; Score 57.8; DB 13; Length 1201;  
 Best Local Similarity 38.0%; Pred. No. 0.32;  
 Matches 167; Conservative 43; Mismatches 230; Indels 0; Gaps 0;

131 TTGCTATTATTCAGGCGTAATGGAGCTCATTTGAAACAATTTCTCTATTTTAA 190  
 |||||  
 1106 TTTTATTTTATTAATATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1047  
 191 CAGAAATCATTTGGTGGTGAAGTACAGCAGTATGCTTATTTATTTATTTGTTG 250  
 1046 TTTATTAATWNAATTTTAAKATATATATATTTTATTTTATTTTATTTTATTTT 987  
 251 CTCAGAGATAGCTATTTTATTTATTTGATGCAATATTTACTTTTATTTGCTTACGCAATGGA 310  
 986 TTTAAATATTTTATTTATTTATTTATTTTAAANMTTATTTTATTTTATTTTATTTT 927  
 311 TTTCAATTCAGTACAGTGGTGCAATATGCACTACAAAGTTTAAATTTAAATGGA 370  
 926 ATATATATTTTATTTAAATATATATATATATATATATATATATATATATATATTTT 867  
 371 GTAATTTAT 430  
 866 WTWNTTTTAT 807  
 431 GACTTTTATAGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 490  
 806 ATATTTTAT 747  
 491 AAGGAGCTTTTCAAACTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 550  
 746 ATATTTTAT 687  
 551 ATATGCTTATATACAGGATA 570  
 686 ATTTTATTTAT 667

RESULT 9 568 bp mRNA linear EST 11-SEP-2002  
 BU497346/c  
 LOCUS PF5870ab62b05.y1 Plasmodium falciparum 3D7 asexual cDNA Plasmodium  
 DEFINITION falciparum 3D7 cDNA 5', mRNA sequence.  
 ACCESSION BU497346  
 VERSION BU497346.1 GI:22793540  
 KEYWORDS EST.  
 SOURCE Plasmodium falciparum 3D7  
 ORGANISM Plasmodium falciparum 3D7  
 REFERENCE 1 (bases 1 to 568)  
 AUTHORS Fukuoka, Y., Chakrabarti, D., Haywood, R., Clifton, S., Pape, D.,  
 Marra, K., Hillier, L., Martin, J., Wylie, T., Dante, M., Theising, B.,  
 Bowers, Y., Gibbons, M., Rutter, E., Bennett, J., Jentes, E., Ronko, I.,  
 Tsagaris, V., Richey, J., Wadkins, J., Kennedy, S., Levinson, D.,  
 Maguire, L., Richey, J., Wadkins, J., Kennedy, S., Levinson, D.,  
 Watson, R., Wilson, R., and Sibley, D.  
 TITLE Unpublished  
 JOURNAL Unpublished  
 COMMENT Contact: L. David Sibley  
 WashU Plasmodium EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@sibley.wustl.edu  
 Library was constructed by Debopam Chakrabarti DNA sequencing by:  
 Washington University Genome Sequencing Center for information on  
 obtaining a clone please contact: L. David Sibley  
 (sibley@bocwin.wustl.edu), Washington University  
 Seg primer: -40UP from Gibco

FEATURES  
 source  
 High quality sequence stop: 425.  
 Location/Qualifiers  
 1..568  
 /organism="Plasmodium falciparum 3D7"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:36329"  
 /lab\_host="DH10B (GeneHog, Invitrogen, Inc.)"  
 /clone\_id="Plasmodium falciparum 3D7 asexual cDNA"  
 /note="Vector: pBluescript SK plus, Site 1: EcoRI, Site 2:  
 XhoI; Library was constructed by Debopam Chakrabarti.  
 Total RNA samples were isolated from mixed stage  
 zygotes by the acidic guanidium-phenol chloroform  
 method. The poly A+ RNA was isolated by the polyAT-Tract  
 mRNA isolation system (Promega, WI) using streptavidin  
 magnetic particles. Directional cDNA libraries were  
 constructed by oligo d(T) priming of poly(A)+ RNA (5mg)  
 into EcoRI and XhoI sites of 1 ZapII vector using the Zap  
 cDNA synthesis kit (Stratagene, CA). The average size of the  
 cDNA inserts in the library was between 1.0 and 1.5kb.  
 Clones were mass excised using the Exassist helper phage  
 (Stratagene), the phagemids were precipitated with PEG  
 8000 and extracted with phenol/chloroform. Phagemid DNA  
 was electroporated into DH10B cells."

BASE COUNT 295 a 64 c 57 g 152 t  
 ORIGIN

Query Match 9.2%; Score 57.4; DB 13; Length 568;  
 Best Local Similarity 46.2%; Pred. No. 0.39;  
 Matches 190; Conservative 0; Mismatches 221; Indels 0; Gaps 0;

139 TTGCTTTGAGTTTACAGCAGTATATATATATATATATATATATATATATATATATAT 258  
 |||||  
 568 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 509  
 259 ATAGCTATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 318  
 508 TTGATTAATACAGTTTTCATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 449  
 319 CGTGTACAGTTTGTGCTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 378  
 448 CTTTAAAGTGTGTTTCTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 389  
 379 AAT 438  
 388 TTATTCATACAGATTTTCTATATATATATATATATATATATATATATATATATATAT 329  
 439 CGTATTTAGTTTCAAGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 498  
 328 ATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 269  
 499 TTTTCAAACTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 558  
 268 TCTGAAGAT 209  
 559 AATACAGAT 609  
 208 TTATTAAT 158

RESULT 10 1201 bp mRNA linear EST 15-MAY-2003  
 BX439779/c  
 LOCUS BX439779 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE014YF05  
 DEFINITION 3-PRIME, mRNA sequence.  
 ACCESSION BX439779  
 VERSION BX439779.1 GI:30771778  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 REFERENCE 1 (bases 1 to 1201)  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.



QY 522 TTACGAAGCTCTTACGATGCGCAGATTATAGCTTAATACAGAAATGATCTGTTT 581  
DB 245 ATCTTAATAATATATTTCTTTCTTAATAAATTTGTTATTAATAATATATATA 186  
QY 582 ATATAGCAATATCTATGACATATTT 609  
DB 185 ATTCTTAATCAATTTTATTAATAATTTAT 158

RESULT 12  
LOCUS CNS017V2  
DEFINITION Drosophila melanogaster genome survey sequence Sp6 end of BAC  
BACN37J10 of DrosBAC library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.  
ACCESSION AL108536  
VERSION AL108536.1 GI:5628840  
KEYWORDS GSS.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1101)  
REFERENCE Genoscope.  
AUTHORS Direct Submission  
TITLE Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :  
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
COMMENT Determination of this BAC-end sequence was carried out as part of a  
collaboration with the European Drosophila Genome Project (EDGP) -  
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC  
library (Dros BAC) was made by Alain Billaud at CEPH (Centre  
d'Etude du Polymorphisme Humain) with funding provided by a MRC  
project grant. The DNA was prepared from embryos by Alain Bucheton  
and Genevieve Payan. It has been constructed in the vector  
pBelobAC11.

FEATURES  
source Location/Qualifiers  
1..1101  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone\_1fb="DrosBAC"  
/clone\_1lb="DrosBAC"  
/plasmid="pBelobAC11"  
/note="end : Sp6"

BASE COUNT 268 a 174 c 114 g 366 t 179 others  
ORIGIN  
Query Match 9.1%; Score 56.8; DB 29; Length 1101;  
Best Local Similarity 38.6%; Pred. No. 0.5;  
Matches 136; Conservative 43; Mismatches 173; Indels 0; Gaps 0;

QY 173 TTTTCACATATATTTTACAGATCATTTGCGTTGAGTTACAGACAGTCATATAG 232  
DB 750 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 809  
QY 233 CTATATATTTATGTTGCTCAAGATAGCTATTTATGATGCCAATATTTACTTTTA 292  
DB 810 WTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 869  
QY 293 TTGCTGTTCAGCGATTTTCATTAAGTGTACAGTTGGTCATTATGACTACAAAG 352  
DB 870 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAT 929  
QY 353 TTTTAAATTTAAATGAGTAATTTATATATAAAGGTGAAGAAGCTTGGT 412  
DB 930 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 989  
QY 413 CTCACAAACACTGTTCGATTTTACGATTTAGTTACGATTAATGTTAGTATG 472  
DB 990 WTTTATATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1049

QY 473 TTCCATATATATTTAATAAGAGATTTTCAACTTTTACCATATATTA 524  
DB 1050 TTTTATTAATAATAATATTTATTTATTTATTTATTTATTTATTTATTTA 1101

RESULT 13  
LOCUS CNS010MP/c  
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC  
BACN04L20 of DrosBAC library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.  
ACCESSION AL099163  
VERSION AL099163.1 GI:5610774  
KEYWORDS GSS.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 734)  
REFERENCE Genoscope.  
AUTHORS Direct Submission  
TITLE Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :  
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
COMMENT Determination of this BAC-end sequence was carried out as part of a  
collaboration with the European Drosophila Genome Project (EDGP) -  
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC  
library (Dros BAC) was made by Alain Billaud at CEPH (Centre  
d'Etude du Polymorphisme Humain) with funding provided by a MRC  
project grant. The DNA was prepared from embryos by Alain Bucheton  
and Genevieve Payan. It has been constructed in the vector  
pBelobAC11.

FEATURES  
source Location/Qualifiers  
1..734  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone\_1fb="DrosBAC"  
/clone\_1lb="DrosBAC"  
/plasmid="pBelobAC11"  
/note="end : T7"

BASE COUNT 288 a 62 c 2 g 211 t 171 others  
ORIGIN  
Query Match 9.1%; Score 56.6; DB 29; Length 734;  
Best Local Similarity 35.5%; Pred. No. 0.55;  
Matches 124; Conservative 65; Mismatches 159; Indels 1; Gaps 1;

QY 178 TACTATATTTTACGAATCATTTGCGTTGAGTTACAGACAGTCATATAGCTTCA 237  
DB 371 TTTTATGATGTTTATTTTATTTATTTATTTATTTTATTTTATTTTATTTT 312  
QY 238 TTTTATGTTGCTCAAGATAGCTATTTATGATGCCAATATTTACTTTTATGCT 297  
DB 311 TTTTATGTTGCTCAAGATAGCTATTTATGATGCCAATATTTACTTTTATGCT 252  
QY 298 GTTACGATGATTTTATTAAGTGTACAGTTGGTCATTATGACTACAAAGGTTT 357  
DB 251 WTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 192  
QY 358 AAATTTAAATGAGTAATTTATATATAAAGGTGAAGAAG-AGTTTGTCTTCA 416  
DB 191 WMAAATAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 132  
QY 417 ACAACACTGTTCGATTTTACGATTTAGTTTACGATTAATGTTATAGTATGTC 476  
DB 131 AATATTAATATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 72  
QY 477 ATATATGATTTAATAAGAGATTTTCAACTTTTACCATATATTAAT 525  
DB 71 ATATAKATATATATATATTTTATTTTATTTTATTTTATTTTATTTTATTTT 23

RESULT 14

CNS000B8

### DEFINITION

CNS00008 1101 bp DNA linear GSS 03-JUN-1999  
Drosophila melanogaster genome survey sequence TET3 end of BAC #  
BACR01A24 of RPI-8 library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.

**ACCESSION**

VERSION

## KEYWORDS

## SOURCE

## ORGANISM

*Drosophila melanogaster* (fruit fly)

*Drosophila melanogaster*

Eukaryota; Metazoa; Art

## REFERENCE

## AUTHORS

TITLE

JOURNAL

**COMMENT**

Submitted (2002-JUN-1999) Genoscope - Centre National de Séquençage  
BP 131 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr  
- Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr))

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley/ *Drosophila* Genome Project (BDGP). The BDGP is constructing a physical map of the *Drosophila* melanogaster genome using these BACs. For further information please see <http://www.flyritely.org> The BDGP *Drosophila* melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Cammocer at the Pleier de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCL-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2. cm bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BAC/PC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

**FEATURES**  
**source**

Location/Qualifiers  
1. .1101

BASE COUNT	228 a	114 c	110 g	512 t	137 others
ORIGIN					

BASE COUNT	228 a	114 c	110 g	512 t	137 others
RIGIN					

QY	167	AAACAATTTTCTACAAATTTTTCACAGATCATTTGCGTTGAGGTTCACAGACAGCTAG	226
Db	682	AAATATTTTTTCTKXGKMTDITTTTGGTWTATTTTGGDAATWTATTTTTTTTWTATATMTM	741
QY	227	TAAATGCTTATTTATTTATGTTGCTCAAGAGATAGCTATTTATGTATGCCATATTTAC	286
Db	742	TWTTWTAAATTTTWTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT	801
QY	287	TTTTTATGCTGTACGGCATGATTTTCATTAACGTACAAGTGTGTCATATAGCACA	346
Db	802	TTTWTATTTWTATATMTTWTMTTTTTTTTTTTTATTTATTTTTTTTTTTTATTTTATTTA	861
QY	347	CAAAAGTTTTTAAATTTTAAATGAGCTAAATTTATATATATAAAGGTTGAAAGGAATGT	406
Db	862	AAAAAATTTAATTTATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTATATATTT	917
QY	407	TTGCTTCCACAACAACCTGTTGCTGACTTTTACAGATTTAGTTCAGAGATTTGTTATAG	466
Db	918	TWATMTWTAAATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTATTTATTTA	977
QY	467	GTAATGTTCCAAATATGATTTATATAAAGAGAGCTTTCAAACTTTTACCATATATTTAG	526

QY

Db

038 MAMAIWATWITTAATAITTTTTTTTTTWWAAIAITTTAAITTTWITWITTTTT 1092

## RESULT 15

BX461824/

## LOCUS

## DEFINITION

BX461824 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone

CS0DF034YA08 5-PRIME, m

**SOURCE**

## ORGANIS

**TITLE** Full-length cDNA libraries and normalization

Mammalia; Eutheria; Primates; Carnivora; Hominoidea; Homo.  
1 (bases 1 to 1201)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact: Genoscope  
Genoscope - Centre National de Séquençage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
was not normalized. Library was constructed by Life Technologies,  
division of Invitrogen. This sequence belongs to sequence cluster  
8170.r For more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgf.bln/cluster.cgi?seq=CS0DP034BA040P1&cluster=8170.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/InvitrogenCorporation1600  
Paradise Avenue Genoscope sequence ID : CS0DP034BA040P1.

**FEATURES**  
**source**

Location/Qualifiers  
1. .1201

BASE COUNT	385 a	157 c	102 g	312 t	245 others
------------	-------	-------	-------	-------	------------

BASE COUNT	385 a	157 c	102 g	312 t	245 others
------------	-------	-------	-------	-------	------------

Query Match	9.1%;	Score 56.6;	DB 13;	Length 1201;
Best Local Similarity	31.7%;	Pred. No. 0.54;		
Matches	152;	Conservative	96;	Mismatches 230; Indels 1; Gaps 1
QY	120	GATGCTGGGGCTTGCTATTATTACGGCGGTATGGGACGTCATTTTGGACAATTTCTA	179	
DB	1152	KWTTTGGKGAKWTKMKAKDGAATTTADADTRRDVDTATTTTWTWAAADMTAWTTT	1093	
QY	180	CTATATTTTTCACAGATCATTTGGTTTGGAGTTACGACAGACAGATATGCTTATTT	239	
DB	1092	TTTTTATWAGHKAPAADMAKAAATWATWTTAAATKTKMAAAAKAATKAKATTTWD-TTTT	103	
QY	240	TATTTATGTTGCTCAAGATAGTATTTTATGATGCCAATATTACTTTTATATGCTGT	299	
DB	1033	WATTAARATATTTWTAATAAAAKATATAATATATKTKGAAAAATTTWATKTKTKAKAKAW	974	
QY	300	TACGGCAGATTTTCATTAAGTGACAAAGTGTCATTTATGACTAACAAAGTTTAA	359	
DB	973	TKCATKATKATTKTKAKRTTTTATGAGWATTTWARTTAAWAGAAATTTTTTTTATATWATA	914	

	AL565455/c	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Qy	360	ATTATAAATGGAGTAAATTTATATATATAAAGGTTGAAGAAATGTTGGCTTCACACA	1201 bp	AL565455	1		Homo sapiens (human)	Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	Full-length cDNA libraries and normalization	unpublished	On Feb 16, 2001 this sequence version replaced gi:12916948.
Db	913	ATWTTTATATWTTWANTTAKAATATTGGATADMDANTTTTTTAATTKATTWMTGATRTK		AL565455	1		Homo sapiens (human)	Homo sapiens	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
Qy	420	AACACTTGTCGACTTTTACGTAGTTAGTTCAGTAATGTTATAGGTAATGTTCCATA		AL565455	1		Homo sapiens (human)	Homo sapiens					
Db	853	WGAKAKMDTFAKAKATWARTTAAABAAAATTGTTWTAATBRATBRBATATATATATTTT		AL565455	1		Homo sapiens (human)	Homo sapiens					
Qy	480	TATGATTATAAAAGAGAGTITTCAACTTTTACCATTAATATATGCAAGTCCTTCAG		AL565455	1		Homo sapiens (human)	Homo sapiens					
Db	793	ATWATWATAWAAWATAKATTTKTTTAAAKRTWAAWATTAAGTARWAAATATATATAGT		AL565455	1		Homo sapiens (human)	Homo sapiens					
Qy	540	TGTGGCAGATTATATAGCTTAATACAGAAATAGTACTGTTTATATACGATATTCCTA		AL565455	1		Homo sapiens (human)	Homo sapiens					
Db	733	KTKTKAAWATKTTWGAATGKTKAAKTATTAKTARTTGKGMAAATTKTGATTATATTTTA		AL565455	1		Homo sapiens (human)	Homo sapiens					

Genoscope - Centre National de Séquençage  
BP 191 91006 Evry cedex - France  
Email: [segrete@genoscope.cns.fr](mailto:segrete@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 9232.f For  
more information about this cluster, see  
<http://www.genoscope.cns.fr/>  
[cg3-bim.cluster.cgi?seq=CS0DP005BH09NP1&cluster=9232.f](http://cg3-bim.cluster.cgi?seq=CS0DP005BH09NP1&cluster=9232.f). Contact :  
Peng Liang Email : [liang@lifetech.com](mailto:liang@lifetech.com) URL :  
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CS0DP005BH09NP1.  
Location/Qualifiers  
1..1201

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODF005X018"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_1b="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

```

Query Match	9.0%;	Score 56.2;	DB 9;	Length 1201;
Best Local Similarity	34.2%;	Pred. No. 0.65;		
Matches	155;	Conservative	76;	Mismatches 219; Indels 3; Gaps 1

  

QY	131	TTGCTATTATTATTCAGCGCGTAATGGAGCGCTATTGGAACAATTTTCTACTATATTTTAA	190
DB	1196	TWTATWTATWTATWTATWTAT	113
QY	191	CAGATCAATTCGGGTTGAGTTAGCAGCAGCAGCATATATCTTATTTATTTAATGGTGG	250

[illegible]

RESULT 17				
LOCUS	EX462546			
DEFINITION	EX462546 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens CDNA clone COSDH003yp10 3-PRIME, mRNA sequence.	1146 bp	linear	EST 22-MAY-2003
ACCESSION	EX462546			
VERSION	EX462546.1	GI:11027451		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Bukhariyoti, Metaroca; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 1146)			
TITLE	L1,W.B., Gruber,C., Jesse,J. and Polayes,D.			
JOURNAL	Full-length cDNA libraries and normalization			
COMMENT	Unpublished			
	Contact: Genoscape			
	Genoscope - Centre National de Sequencage			

Email: [segrete@genoscope.cns.fr](mailto:segrete@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 994.f For  
more information about this cluster, see  
[http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DH003DH05NP1&c1uster=994.f](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DH003DH05NP1&c1uster=994.f). Contact :  
Feng Liang Email : [liang@lifetech.com](mailto:liang@lifetech.com) URL :  
<http://fulllength.invitrogen.com>, Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DH003DH05NP1.

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DH003YP10"
/tissue_type="T CELLS (JURKAT CELL LINE)"
/cell_line="JURKAT CELL LINE"
/clone_id="Homo sapiens T CELLS (JURKAT CELL LINE)"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
BASE COUNT      216 a      127 c      154 g      436 t      213 others
ORIGIN
Query Match      9.0%, Score 56; DB 13; Length 1146;

```



Best Local Similarity 35.8%; Pred. No. 0.71;  
Matches 163; Conservative 67; Mismatches 223; Indels 2; Gaps 1;

QY 162 TTTTGAACAATTTCTACTATATTTTACAGAAATCATTCGCTGAGTTACAGCA 221  
DB 645 TTTTATATWATTTATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT  
QY 222 GTCAATATATGCTTATTTATTTATTTATGCTCAAGATAGCTATTTT -ATGATGCCA 279  
DB 705 TTTATATWATTTTDRTTTTTTTTTTTTTTKTRTGAATAATDTTTTAAATTTTTTTTK 764  
QY 280 ATATATCTTTTATGCTGATCGCAATTCATTTACGTTCAAGTGTGATTA 339  
DB 765 GTGGTTTTTAAATWRTTTTTTTTATTTTAAATTAADWTCKGRAAATWTWWT 824  
QY 340 TGCACTCAAAAGTTTAAATTTAAATGAGTAAATTAATATATTAATTAAGCGTTGAA 399  
DB 825 TTTATTTTAAAAAATTTAAATWTDGATTTTGTGTTTTTTTAAATTTATTDAR 884  
QY 400 GGAATGTTGCTTCAACAAACACTTGTGACTTTTACGATTTAGTTCAAGTAAT 459  
DB 885 AAGKTTWAMRTWTAGATAAAAAATTTTTRTTKTTKTTTAAABAMRTAATTT 944  
QY 460 GTTATAGTATTTGTCATATATGATTTAAAGAGAGTTTCAACTTTTACATTA 519  
DB 945 KTKTAAATAATTTTTTTTAAARKMAAMDAAATDWAATTTWAAAAAATWTTTGT 1004  
QY 520 TATATGCAAGTCCTGAGTGTGAGATATATATGCTATATACGAAATGACTGTT 579  
DB 1005 TTTTATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1064  
QY 580 TTAATATAGCTAATTCCTATGACATTTATTCAGT 614  
DB 1065 TTTTATTAATAWMTTAAATTTTAAARATTTADT 1099

RESULT 18  
LOCUS CNS003DQ 1101 bp DNA linear GSS 03-JUN-1999  
DEFINITION Drosophila melanogaster genome survey sequence TEXT end of BAC #  
BAC080109 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
ACCESSION AL064580  
VERSION AL064580.1 GI:4941932  
KEYWORDS GSS.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 1101)  
AUTHORS Genoscope.  
TITLES Direct Submission  
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqef@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Oseogawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).  
FEATURES  
Source location/Qualifiers  
1..1101

/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BAC080109"  
/clone\_1db="RPCI-98"  
/note="end : TET3"

BASE COUNT 291 a 51 c 117 g 404 t 238 others  
ORIGIN

Query Match 9.0%; Score 55.8; DB 29; Length 1101;  
Best Local Similarity 41.6%; Pred. No. 0.77; Indels 3; Gaps 1;  
Matches 185; Conservative 23; Mismatches 234;

QY 131 TTGCTATTTATTCAGCGGTATAGGAGTCAATTTTGAACAATTTCTACTATATTTTA 190  
DB 18 TTTTNNNNNNNNNNNNNNNNNNNGTGGNNNNNTTTTAAATTAATTBAMWTTTTTTTT 77  
QY 191 CAGATCATTTGCGTTGAGGTACAGACAGTCAGATATGCTTATTTATATGTTG 250  
DB 78 TTTATATTAAABA 137  
QY 251 CTCAGAGATGCTATTTTATGATGCCAATATCTTTTATGCTTACGCAATGA 310  
DB 138 CGAAAAATWATATTT 197  
QY 311 TTTCTATACGTCAAGATTTGTCATTTATGACATTAAGATTTTAAATTTAAATGA 370  
DB 198 TTTTATTTTATTTTAAATTTTAAATTTTATTTTATTTTATTTTATTTTATTTT 255  
QY 371 GTAATTTATATATATTAAGGTTGAAGAGATGTTGCTTCAACAAACCTGTTC 430  
DB 256 -TTTWTGTAATTAATTTTAAATWATATWAAATWATTAATTAATTAATTAATTA 314  
QY 431 GACTTACAGTATTTGTTCAAGTATGTTAAGATTTGTTCAATATATGATTAAT 490  
DB 315 AATATATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 374  
QY 491 AAGGAGATTTTCAACCTTTTACATTTATTTATGCAAGCTTCAGTGTGACAGAT 550  
DB 375 WATTTTATTTTAAATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTT 434  
QY 551 ATATGCTTAATACGAATGACT 575  
DB 435 WTWTATTTTAAATTAATTTATTTWATT 459

RESULT 19  
LOCUS BX415819 977 bp mRNA linear EST 13-MAY-2003  
DEFINITION BX415819 Homo sapiens THYMUS Homo sapiens cDNA clone CSOCAP008YG19  
5-PRIME, mRNA sequence.  
ACCESSION BX415819  
VERSION BX415819.1 GI:30645987  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 977)  
AUTHORS L.W.B., Gruber,C., Jesse,J. and Polayes,D.  
TITLES Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqef@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@life-tech.com  
<http://fulllength.invitrogen.com/invitrogen> 1600  
Faraday Avenue genoscope sequence ID : CSOCAP008AD10QPL.  
FEATURES  
Source location/Qualifiers  
1..977  
/organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="COSCAP008Y19"
/tissue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

```

BASE COUNT 240 a 144 c 98 g 380 t 115 others

Query Match 8.9%; Score 55.4; DB 13; Length 977;

Best Local Similarity 37.0%; Pred. No. 0.93;

Matches 141; Conservative 51; Mismatches 189; Indels 0; Gaps 0;

```

145 GGGCTAATGGAGCGTCATTGGAACAATTTCTACTAATTTTACAGAAATCATTTGG 204
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
355 KGGKGRGDRDDKDKKKKKAAAAAADDKDKTKTTTWTWTWTWTWTWTWTWTWT 414
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
205 TTGAGGTACAGACAGTCAGTATAGCTTATTTATTTATGTCACAGATAGCT 264
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
415 TKTITWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 474
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
265 ATTTTATGATGCAATATTTACTTTTATGCTGTACGCGCATGATTCATACGTGA 324
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
475 TTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 534
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
325 CAAGTGGCATATGACATCAAAAGTTTAAATTAATGAGTAATTAATA 384
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
535 TATTTTGTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 594
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
385 ATAAAGGTTGAAGAAATGTTGCTTCACAAACAATTGTCAGCTTACGTAGT 444
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
595 WTCTTAAAAAATTTAAATTTTAAAGDTAAATWTAAATTTTWTWTWTWTWT 654
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
445 TTAGTCAAGTAATGTTATAGTATTTGCTCAATATGATTTAAAGAGAGTTTCA 504
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
655 TMAAATTAACCTAAATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 714
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
505 AACTTTTACCATTAATTAAT 525
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
715 AAWTTTTTAAAGAAATTT 735
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

## RESULT 20

BO452226 473 bp mRNA linear EST 29-MAY-2002

LOCUS PESTOa33908.y1 Plasmodium falciparum 3D7 asexual cDNA plasmodium

DEFINITION falciparum 3D7 cDNA 5' mRNA sequence.

ACCESSION BO452226

VERSION BO452226.1 GI:21255338

KEYWORDS EST

SOURCE Plasmodium falciparum 3D7

ORGANISM Plasmodium falciparum 3D7

REFERENCE 1 (bases 1 to 473)

AUTHORS Tang, K., Cole, R., Chakrabarti, D., Haywood, R., Clifton, S., Page, D.,

Matta, M., Hillier, L., Martin, J., Wylie, T., Dante, M., Theising, B.,

Bowers, Y., Gibbons, M., Ritzer, E., Bennett, J., Jentes, E., Ronko, I.,

Tagareishvili, R., Belaygorod, L., Franklin, C., Carr, L., Grow, A.,

Maguire, L., Richey, J., Wadkins, J., Kennedy, S., Levinso, D.,

Waterston, R., Wilson, R., and Sibley, D.

Washu Plasmodium EST Project

Unpublished

CONTACT: L. David Sibley

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.wustl.edu

## FEATURES

## SOURCE

Library was constructed by Debopam Chakrabarti DNA sequencing by:  
Washington University Genome Sequencing Center for information on  
obtaining a clone please contact: L. David Sibley  
(sibley@orcim.wustl.edu), Washington University  
Seq primer: -400P from Gibco  
High quality sequence stop: 354.  
location/Qualifiers

1. 473

/organism="Plasmodium falciparum 3D7"

/mol\_type="mRNA"

/db\_xref="taxon:36329"

/lab\_host="DH10B (GeneHog, Invitrogen, Inc.)"

/clone\_lib="Plasmodium falciparum 3D7 asexual cDNA"

/notes="Vector: pBluescript SK plus; Site 1: EcoRI; Site 2:

XhoI; Library was constructed by Debopam Chakrabarti.

Total RNA samples were isolated from mixed stage

sapronin(0.1%) lysed P. falciparum 3D7 infected

erythrocytes by the acidic guanidium-phenol chloroform

method. The poly A+ RNA was isolated by the polyAT-tract

mRNA isolation system (Promega, WI) using streptavidin

Magnosphere particles. Directional cDNA libraries were

constructed by oligo d(T) priming of poly(A)+ RNA (5mg)

into EcoRI and XhoI sites of 1 ZapII vector using the Zap

cDNA synthesis kit (Stratagene, CA). The average size of the

cDNA inserts in the library was between 1.0 and 1.5kb.

Clones were mass excised using the ExAssist helper phage

(Stratagene), the phagmids were preclipped with PBG

8000 and extracted into DH10B cells."

was electroporated into DH10B cells."

## BASE COUNT

129 a 48 c 52 g 244 t

## ORIGIN

Query Match 8.9%; Score 55.2; DB 13; Length 473;

Best Local Similarity 46.4%; Pred. No. 1;

Matches 180; Conservative 0; Mismatches 208; Indels 0; Gaps 0;

```

222 GTCACTATATGCTTATTTATTTATTTATTTGCTCAAGATACCTATTTATGATGCCAT 281
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2   GTCACTTGAATTTTGTGCTTTTCTACAGCTAATTTTGATTAACAGTTTGTGCTT 61
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
282 ATTACTTTTATTTGCTGTTAGCGCATGATTTCAATAGCTGACAAAGTTGGCATATG 341
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 TATATTTTGAATTTGTTATCTTTTGTTCACACTCTTTTAAAGTGTGTTCTTATTT 121
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
342 GACTCAAAAGTTTAAATTTAAATGAGTAATTTAATTAATTAATTAATTAATTAAT 401
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
122 ATTAAATTTGCTTTATTTAGTTGATTAATTTATTTGTTATTTCAATCAAGATTAAT 181
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
402 AATGTTTGTCTCTCAACAAACAATTGTCAGCTTTACGTAGTTAGTTCAAGTAATGT 461
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
182 ATGACTAGATATGCTTAATTTATTTGTTGAATTTTATTTTATTTTAAATTAATCT 241
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
462 TATAGTATTTGCTCAATATGATTAATTAAGAGAGTTTCAACTTTTACCAATTA 521
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
242 TATGAGTATTTGCTTATCTGAATCAATGAGTATGAGTATTAATTAATTAATTAAT 301
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
522 TATGCAAGTCTTCAAGTGTGAGATTAATTAATGCTTAATCAAGAAATGATCTGTTT 581
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
302 ATCTTAAATTAATTTTCTTTTCAATTAATTTTGGTATTAATTAATTAATTAATTA 361
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
582 ATATACGTAATTTCTTATGACATTAATTT 609
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
362 ATTCCTATCATTTTATTAATTTAATTTAAT 389
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

## RESULT 21

CNS04DOK/c

LOCUS CNS04DOK/c

DEFINITION Tetradon nigroviridis genome survey sequence 17 end of clone

sequence.

ACCESSION AL285149

VERSION AL285149.1 GI:8023560

945 bp DNA linear GSS 01-SEP-2000

Tetradon nigroviridis genome survey sequence 17 end of clone

101H21 of library G from Tetradon nigroviridis, genomic survey

sequence.

AL285149

AL285149.1 GI:8023560

KEYWORDS GSS; genome survey sequence.

SOURCE Tetracodon nigroviridis

ORGANISM Tetracodon nigroviridis

REFERENCE 1

AUTHORS Roest Crolius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fzames, C., Wincker, P., Broctier, P., Quetier, F., Saurin, W. and Weissenbach, J.

TITLE Estimate of human gene number provided by genome-wide analysis using Tetracodon nigroviridis DNA sequence

JOURNAL Nat. Genet. 25 (2), 235-238 (2000)

MEDLINE 20296633

PUBMED 10835645

REFERENCE 2

AUTHORS Roest Crolius, H., Jallion, O., Dasilva, C., Ozouf-Costaz, C., Fzames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.

TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetracodon nigroviridis

JOURNAL Genome Res. 10 (7), 939-949 (2000)

MEDLINE 20359837

PUBMED 10899143

REFERENCE 3

AUTHORS (bases 1 to 945)

TITLE Genoscope.

JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)

COMMENT This sequence is a single read and was generated as part of a large scale clone and sequencing project of the Tetracodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/tetracodon>.

FEATURES

source

1. .945

/organism="Tetracodon nigroviridis"

/mol\_type="Genomic DNA"

/db\_xref="taxon:99863"

/clone="101H21"

/clone\_1lb="G"

/note="Genoscope sequence ID : COBG101CD11LP1-end : T7"

BASE COUNT 386 a 112 c 96 g 231 t 120 others

ORIGIN

Query Match 8.8%; Score 54.8; DB 29; Length 945;

Best Local Similarity 37.2%; Pred. No. 1.2;

Matches 172; Conservative 52; Mismatches 237; Indels 1; Gaps 1;

Oy 159 TCATTGGAACAATTTCTACTATTTTACAGAAATCATTTGGCTTGAAGTTACAGC 218

Db 902 TTATTTTATATATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAN 843

Oy 219 ACAGCTATATATGCTTATTTATTTATTTATTTGCTCAAGATAGCTATTTATGATGCC 278

Db 842 WTAAWTTATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 784

Oy 279 AATATTTACTTTTATTTGCTTTACGCGATGATTTTCAATTCAGTACAGTTGGTGCAAT 338

Db 783 TTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 724

Oy 339 ATGACATCAAGAGTTTAAATTTAAATGAGTAATTTAAATTTAAATGAGTTGAA 398

Db 723 TTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 664

Oy 399 AGGATGTTGCTTCTCAACAACACTTTGCACTTTGAGTTTATGTTAGTTCAAGTAT 458

Db 663 WTTWTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 604

Oy 459 TGTATATAGTATTTGTTCCATATATGATATTAAGAGAGTTTCAACCTTTTACCAAT 518

Db 603 ATTWTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 544

Oy 519 ATATTATGACAGTCCCTCAGGTGCGAGATTATATGCTTTATATACAGAAATAGTACTGT 578

Db 543 ATWTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 484

Oy 579 TTATTTATGCTTAATTTCTATGACAAATTTATTTGAGTGCAGCA 620

Db 483 ATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 442

RESULT 22

AL536104 1201 bp mRNA linear EST 31-MAY-2003

LOCUS

DEFINITION AL536104 Homo sapiens PBTAL BRAIN Homo sapiens cDNA clone

ACCESSION CS0DF022YC18 5-PRIME, mRNA sequence.

VERSION

AL536104

KEYWORDS

SOURCE

ORGANISM

Al536104.2 GI:31260974

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1

AUTHORS Li, W. B., Gruber, C., Jesse, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT On Feb 13, 2001 this sequence version replaced gi:1279597.

Contact: Genoscope - Centre National de Sequencage

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: sequef@genoscope.cns.fr

Library was constructed by Life Technologies, a division of Invitrogen. Contact: Feng Liang Email: fliang@life.com URL: <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Paraday Avenue Genoscope sequence ID : CS0DF022BB09QPI.

FEATURES

source

1. .1201

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DF022YC18"

/tissue\_type="FETAL BRAIN"

/dev\_stage="fetal"

/clone\_1lb="Homo sapiens PBTAL BRAIN"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

BASE COUNT 359 a 119 c 146 g 409 t 168 others

ORIGIN

Query Match 8.8%; Score 54.8; DB 9; Length 1201;

Best Local Similarity 29.6%; Pred. No. 1.2;

Matches 126; Conservative 92; Mismatches 208; Indels 0; Gaps 0;

Oy 159 TCATTGGAACAATTTCTACTATTTTACAGAAATCATTTGGCTTGAAGTTACAGC 218

Db 776 TTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 835

Oy 219 ACAGCTATATGCTTATTTATTTATTTATTTGCTCAAGATAGCTATTTATGATGCC 278

Db 836 AATAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 895

Oy 279 AATTTATCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 338

Db 896 TTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 955

Oy 339 ATGACATCAAGAGTTTAAATTTAAATGAGTAATTTAAATTTAAATGAGTTGAA 398

Db 956 AAGWRTTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1015

Oy 399 AGGATGTTGCTTCTCAACAACACTTTGCACTTTGAGTTTATGTTAGTTCAAGTAT 458

QY	Db	1016	TTTWTTTTWTATWSGTATATATAAAAAAATAMTWTMTTTTTTTTAACTMTAMATAMTWTATAAAMWAM	1075
QY	459	TGTTATAGATGTATGTCCATATATGATATATAAAGAGAGATTTCAACTTTTACAT	518	
Db	1076	WRTMTTWTATARAAMAMWAMWTATTTTWTATWMAAAAAAATAADTAATAMTWTGAT	1135	
QY	519	ATATTATGCAAGTCCTTCAGGTGTGGCAGATTAATATGCTATATACAGAAATAGACTTGT	578	
Db	1136	ATTATTAATAMTWTMTADAMDATTTTWTATTTATAMMTTATCTKTATATATATAATAAATA	1195	
QY	579	TTTATA 584		
Db	1196	WADRWW 1201		
RESULT 23				
LOCUS	CNS00EVL/c	1101 bp	DNA	linear
DEFINITION	CNS00EVL	Drosophila melanogaster genome survey sequence T7 end of BAC:		
	BACR29B23 of RPCT-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.			
ACCESSION	AL069706			
VERSION	AL069706.1	GI:4949849		
KEYWORDS	GSS.			
SOURCE	Drosophila melanogaster (fruit fly)			
ORGANISM	Drosophila melanogaster			
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
REFERENCE	1 (bases 1 to 1101)			
AUTHORS	Genoscope.			
TITLE	Direct Submission			
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequence :			
	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr			
	- Web : www.genoscope.cns.fr)			
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <a href="http://www.fruitfly.org">http://www.fruitfly.org</a> The BDGP Drosophila melanogaster BAC library was prepared by Kazuoto Oseogawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCT-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <a href="http://bacpac.med.buffalo.edu/drosophila_bac.htm">http://bacpac.med.buffalo.edu/drosophila_bac.htm</a> .			
FEATURES	Location/Qualifiers			
source	1..1101			
	/organism="Drosophila melanogaster"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:7227"			
	/clone="BACR29B23"			
	/clone_11b="RPCT-98"			
	/note="end : T7"			
BASE COUNT	419 a 91 c 60 g 299 t 232 others			
ORIGIN				
Query Match	8.8%; Score 54.6; DB 29; Length 1101;			
Best Local Similarity	33.9%; Pred. No. 1.3; Mismatches 182; Indels 1; Gaps 1;			
Matches 131; Conservative 73;				
QY	223	TCAGTATATGCTTATTTATTTATGTCCTCAAGATAGCTATTTTATGATGCAATA	282	
Db	1101	TTTMTATTTTMTTTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMTT	1042	
QY	283	TTACTTTTATTCGCTTACGCGCATGATTCATTACGTGTCATGTCGTCATATGCG	342	
Db	1041	TWATATATATATTTTATAAT	982	

[illegible]

Db	684	MMMWATWMTATTTTTTYYUUTTHNNHYTMWNHMHNCSTGHMWTYTMHCSTTHNYCC	743
Oy	287	TTTTATTCGCTGTACGGCATGCATTCAITACSGTGACAGTTGGTGCATTAGSACTA	346
Db	744	YCMHHVTTTTTYUHNHTYTWMNHAHNTTWTTHTTHNMWWMMNHHTTWMTTYW	803
Oy	347	CAGAAGTTTTTAATTTGAATGAGTAATTTAAATATAATAAAGGTGAAGAAGT	406
Db	804	TTTTWTYTTTTWHHTHTHTAHHTWMTTACTHWAMTMAAAAAMAMTNNHTHTWAT	863
Oy	407	TTCCTTCAACAACAACTGGTGCATTTTAACTAGTTAGTTAGTCAAGTAATGTATAG	466
Db	864	MHMVAMTVAHMAHMAWTWTWTWTHHTWMTATNNHMHNCSTHTTHMMMAAMAAM	923
Oy	467	GTAATGTTCCATATATGATATATAAAGAGAGTTTCAAACCTTTTACCATTATATAG	526
Db	924	HTMMMYMTTWTWAWHMYMTWTWHTHTWHTYTTCTWMYHYTTTMMHTTWTHMMYC	983
Oy	527	CAAGTCT 534	
Db	984	HMHMMH 991	
RESULT 25			
CNS020K7/c			
LOCUS			
DEFINITION	CNS020K7	1092 bp DNA linear	GSS 01-SEP-2000
ACCESSION		Tetradodon nigroviridis genome survey sequence T7 end of clone	
VERSION		222L11 of library G from Tetradodon nigroviridis, genomic survey	
KEYWORDS		sequence.	
SOURCE	AL175696.1	GI:7813753	
ORGANISM	GSS; genome survey sequence.		
	Tetradodon nigroviridis		
	Tetradodon nigroviridis		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;		
	Tetraodontiformes; Tetraodontidae; Tetraodontoideae; Tetraodon.		
REFERENCE	1	Roeest Crolius,H., Jalllon,O., Desilva,C., Bouneau,L., Fisher,C., Bernot,A., Fitzames,C., Winkler,P., Brotlier,P., Quetier,F., Saurin,W. and Weissenbach,J.	
AUTHORS		Estimate of human gene number provided by genome-wide analysis using Tetradodon nigroviridis DNA sequence	
TITLE	JOURNAL MEDLINE	Nat. Genet. 25 (2), 235-238 (2000)	
JOURNAL	PUBMED	20296633	
REFERENCE	PUBMED	10835645	
AUTHORS	2	Roeest Crolius,H., Jalllon,O., Desilva,C., Ozouf-Costaz,C., Fitzames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.	
TITLE		Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradodon nigroviridis	
JOURNAL	Genome Res.	10 (7), 939-949 (2000)	
MEDLINE	20359837		
PUBMED	10899143		
REFERENCE	3	(bases 1 to 1092)	
AUTHORS		Genoscope.	
TITLE		Direct Submission	
JOURNAL		Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr	
COMMENT		- Web : www.genoscope.cns.fr)	
		This sequence is a single read and was generated as part of a large scale clone-and-sequencing project of the Tetradodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetradodon.	
FEATURES		Location/Qualifiers	
Source		1..1092	
		/organism="Tetradodon nigroviridis"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:99883"	
		/clone="222L11"	
		/clone_lib="g"	

```

BASE COUNT      383 a      169 c      165 g      262 t      113 others
ORIGIN

Query Match          8.7% ; Score 54.4 ; DB 29 ; Length 1092 ;
Best Local Similarity 37.4% ; Pred. No. 1.4 ;
Matches 142; Conservative 45; Mismatches 193; Indels 0; Gaps 0

OY    230   ATGCTTTATTAATTAGTGTGCCTCAAGAGTAGCATTTTATGTAGVCGAATTACTTT 289
       :|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    1060   WTTTTTAATTTATTTTATTTTATTTTATATAATAAAATTTTAATTWTATTTATTTATTT 1001
OY    280   TTATTCGCTGACGCCAGCATTCATTACGTGACGTACAAGTGTCATTAGACTACACA 349
       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    1000   TTTTTTTTTTTTAAAAAAAATTAAMWTTTNNMTTATATTTTTTTTTTTTTTTTAAAMA 941
OY    350   AGGTTTTTAATTTAAATGAGTAATTTATATATATAAAGGGTTGAAGGATGTTTG 409
       ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    940   AAAWWAAAAAAAAAAAAAAAAATTNAAWTTWAAMAAAAAAMWMTTWTTTTAAATTTT 881
OY    410   CTTCACAACAACACTGTGTGCACCTTTACGTAGTTAGTTCAAGTAATGTTATAGSTA 469
       ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    880   TTTTAAATTTAAATTTTAAATTTTAAATTTTAAAMATATATTTWTTTAAWTAATAA 821
OY    470   TTGTTCCATATATGATTATATAAAGAGAGCTTTCAAACCTTTTACCATTATATTGCA 529
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    820   WTNTTTTWTATAAATAWMAAWMTATWMTTAAAMWTAAMWMTTWTTWTTTTWWWMTW 761
OY    530   GTCTTCAGGTGTGGCAGATTATATGCTTAATACAGAAATGATCTGTTTATATAGC 589
       |:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    760   AAAAAAAMAAWMTWMTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 701
OY    590   TAATTCCTATGACATTAATT 609
       |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    700   AAAAAAAAAAAAAAAAAATTTT 681

RESULT 26
CN5014DY      1021 bp     DNA        linear     GSS 26-JUL-1999
LOCUS         Drosophila melanogaster genome survey sequence T7 end of BAC
DEFINITION   BACNI1004 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION    AL104032
VERSION      AL104032.1 GI:5615643
KEYWORDS     GSS.
SOURCE       Drosophila melanogaster (fruit fly)
ORGANISM     Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE    1 (bases 1 to 1021)
AUTHORS      Genoscope.
TITLE        Direct Submission
JOURNAL      Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT      - Web : www.genoscope.cns.fr
              Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billard at CBPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.

FEATURES             Location/Qualifiers
     source            1..1021
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACNI1004"
                     /clone_lib="DrosBAC"
                     /plasmid="pBelobAC11"

```

BASE COUNT 342 a /note="end : T7"  
 ORIGIN 1 c 104 g 462 t 112 others

Query Match 8.7%; Score 54.2; DB 29; Length 1021;  
 Best Local Similarity 38.2%; Pred. No. 1.6;  
 Matches 105; Conservative 44; Mismatches 126; Indels 0; Gaps 0;

QY 172 ATTTCCTACATATTTTACAGATCATTGCGTTGAGGTACAGACAGTCATATAT 231

DB 734 AKKGGGTGTTAAATATATATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 793

QY 232 GCTTTATTTATTTATGTTGCTCAGAGATGATGATTTTATTTATTTATTTATTT 291

DB 794 TTTTATTTATTTTGGGTTTATTTGATTTAATTTGGGATGATATATATATATAT 853

QY 292 ATTGCTGTACGCGATGATTTTATTTACGTTGACGTTGATTTGATTTGATTTA 351

DB 854 TTTATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 913

QY 352 GTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 411

DB 914 TATATATATATTTTGTGAAATGAAATTTTAAATATTTTAAATATTTTATTT 973

QY 412 TCTCAACAACACTTGTGACTTTTACGTTGTT 446

DB 974 TAAAAAAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1008

RESULT 27  
 AL522840 854 bp mRNA linear EST 22-MAY-2003  
 LOCUS AL522840 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens

DEFINITION CDNA clone CS0DB009YH19 5-PRIME, mRNA sequence.  
 ACCESSION AL522840  
 VERSION AL522840.2 GI:31041104

KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 854)

AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.  
 TITLE Full-length CDNA libraries and normalization

JOURNAL Unpublished  
 COMMENT On Feb 13, 2001 this sequence version replaced gi:12786333.

CONTACT: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France

EMAIL: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 9464.r For  
 more information about this cluster, see

http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0DB009CD10QPLcluster=9464.r. Contact :  
 Feng Liang Email : fliang@life.techn.com URL :  
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Paradey Avenue Genoscope sequence ID : CS0DB009CD10QPL.  
 Location/Qualifiers

FEATURES  
 1..854

source /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DB009YH19"

/tissue\_type="NEUROBLASTOMA COT 10-NORMALIZED"  
 /clone\_id="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"  
 /note="1st strand CDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand CDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 233 a 122 c 114 g 280 t 105 others

ORIGIN

Query Match 8.6%; Score 53.8; DB 9; Length 854;  
 Best Local Similarity 35.2%; Pred. No. 1.9;  
 Matches 148; Conservative 50; Mismatches 223; Indels 0; Gaps 0;

QY 189 TACAGATCATTTCCGTTGAGTTACAGACAGTCAGTATATGCTTATTTATTTAT 248

DB 361 TCCAAAMCCCCBBGTTKKKKKKTTACCTTTTTTTTTTTTTTTTTTTTTTTT 420

QY 249 TCGTCAAGAGATGCTATTTTATTTGATGCGCAATATTTACTTTTATTTGCTGT 308

DB 421 TGGKTATGTTNNNGGKTTTANNNTANCCNNCCNNNNNTTTTATTTATATAT 480

QY 309 GATTTTCATTCAGTACAGTATTTGCTGATTTGACGCTCAAGGTTTATTTAT 368

DB 481 AAAATATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 540

QY 369 GAGTAAATTTATATATATTTAAAGGTTGAAAGAAATGTTGCTTCAACAA 428

DB 541 TTAATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 600

QY 429 TCGACTTTTACGATTTAGTTCAAGTATTTGTTATTTGTTCCATATATGAT 488

DB 601 WATATTTTAAATATTTAAATTTATTTATTTATTTATTTATTTATTTATTT 660

QY 489 AAAAGAGAGTTTCAACTTTTACATTTATTTATTTATTTATTTATTTATTT 548

DB 661 AATATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 720

QY 549 TTAATGCTTATATCAGAAATGATCTTTTATTTATTTATTTATTTATTTAT 608

DB 721 AATATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 780

QY 609 T 609

DB 781 T 781

RESULT 28  
 BX338020 1076 bp mRNA linear EST 02-MAY-2003  
 LOCUS BX338020 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA

DEFINITION clone CS0D1054Y04 5-PRIME, mRNA sequence.  
 ACCESSION BX338020  
 VERSION BX338020.1 GI:30339670

KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1076)

AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.  
 TITLE Full-length CDNA libraries and normalization

JOURNAL Unpublished  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France

EMAIL: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 3701.f For  
 more information about this cluster, see

http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0D1054BA02QPLcluster=3701.f. Contact :  
 Feng Liang Email : fliang@life.techn.com URL :  
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Paradey Avenue Genoscope sequence ID : CS0D1054BA02QPL.  
 Location/Qualifiers

FEATURES  
 1..1076

source /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1054Y04"

/tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_id="Homo sapiens PLACENTA COT 25-NORMALIZED"

BASE COUNT 233 a 122 c 114 g 280 t 105 others

ORIGIN



BASE COUNT 251 a 133 c 28 g 528 t 161 others

/clone="BACN12C18"  
/clone\_lib="drosBAC"  
/plasmid="pbeloBAC11"  
/note="end : 17"

Query Match 8.6%; Score 53.6; DB 29; Length 1101;  
Best Local Similarity 40.7%; Pred. No. 2;  
Matches 179; Conservative 34; Mismatches 226; Indels 1; Gaps 1;

OY 161 ATTGGAACATTTCTCATATTTTACAGAAATCGGTTGAGTACAGC 220  
DB 45 ATTATTAATTTTTCATATTTTATCTTTTATATATATATATTTT 104  
OY 221 AGTCAGTATATCTTATTTATTTATGTCAGAGATAGCTATTTATGATCCAA 280  
DB 105 AATAAATATATATATTTTATTTTATCAATTTATATCTTTTCTTCA 164  
OY 281 TATTAATTTTATGCTGTT-ACGCGATGATTTCTTACGTTACAGATGTCATTA 339  
DB 165 TATTCATATTTTATTTTCTCAAAATTAATATATATTTTATTTT 224  
OY 340 TCGACATCAAGGTTTAAATTAAGAGTAAATTAATAAAGGTTGAA 399  
DB 225 TTTTATATTTATATATATATATATTTTATTTTATTTTATTTT 284  
OY 400 GGAATGTTTCTTCACAAACACTTGTGACATTTTACGATTTAGTCAAGTAT 459  
DB 285 TAAATTTTTCATATATTTTATCTATTTTATTTTATTTTATTTT 344  
OY 460 GTTATAGTATTTGTCATATATGATTAATAAGAGATTTCAACTTTTACATTA 519  
DB 345 TTAATTTTATTTTAAATTTTATTTTATTTTATTTTATTTTATTT 404  
OY 520 TATTTAGCAAGCTTCAGGTCAGATTTATGCTTAATATACAGAAATGACTGTT 579  
DB 405 AAAATCTAAATTTTATTTTATTTTCTTTATTTATTTATTTTATTT 464  
OY 580 TTATATAGCTAATTTCTAT 599  
DB 465 ATCTTATTTTATTTTATTTT 484

## RESULT 31

LOCUS BX441520 887 bp mRNA linear EST 15-MAY-2003

DEFINITION BX441520 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone

ACCESSION CS0DF018YJ11 3-PRIME, mRNA sequence.

VERSION BX441520 GI:30787976

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 887)

Li, W.B., Gruber, C., Jesse, J., and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 7316.r

Contact: Feng Liang Email: filiang@lifetech.com URL: http://fulllength.invitrogen.com/Invitrogen Corporation 1600

## FEATURES

source

1. .887

/organism="Homo sapiens"

/mol\_type="mRNA"

/db xref="taxon:9606"  
/clone="CS0DF018YJ11"  
/tissue\_type="FETAL BRAIN"  
/dev\_stage="fetal"  
/clone\_lib="Homo sapiens FETAL BRAIN"  
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

BASE COUNT 253 a 82 c 94 g 293 t 165 others

Query Match 8.6%; Score 53.4; DB 13; Length 887;  
Best Local Similarity 38.0%; Pred. No. 2.2;  
Matches 141; Conservative 41; Mismatches 189; Indels 0; Gaps 0;

OY 153 GGGACGTCATTTTGAACATTTTCTATATTTTACAGAAATCGTTGAGT 212  
DB 415 GGGCGNNGGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 474  
OY 213 TACAGACAGTCAGTATATGCTTATTTATTTATGTCCTCAAGATAGCTATTAT 272  
DB 475 KKKKKKTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 534  
OY 273 GATGCAATTTACTTTTATTTGCTGTCAGGATGATTCATTAAGTACAAGTTGG 332  
DB 535 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 594  
OY 333 TGCATTTATGATCAACAAAGTTTAAATTTAAATGAGTAAATTAATATAAAG 392  
DB 595 AAAT 654  
OY 393 GTGAAGAAATGTTGCTTCTCAACAAACCTTGTGACTTTTACGATGTTAGTTCA 452  
DB 655 AAAT 714  
OY 453 AGTATTTGTTATGATTTTCTCATATATGATTTTAAAGAGATTTTCAACTTTT 512  
DB 715 WTTTWTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 774  
OY 513 ACCATTATATT 523  
DB 775 WWAATAATTTT 785

## RESULT 32

LOCUS BX357585 1201 bp mRNA linear EST 05-MAY-2003

DEFINITION BX357585 Homo sapiens PLACENTA CON 25-NORMALIZED Homo sapiens cDNA

ACCESSION CS0DI028Y001 5-PRIME, mRNA sequence.

VERSION BX357585 GI:30382176

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1201)

Li, W.B., Gruber, C., Jesse, J., and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 8737.r. For more information about this cluster, see

http://www.genoscope.cns.fr/

cgl-bin/cluster.cgi?seq=CS0DI028AH010P1&cluster=8737.r. Contact :

Feng Liang Email: filiang@lifetech.com URL: http://fulllength.invitrogen.com/Invitrogen Corporation 1600









[illegible]

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
BACR29P01 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	AI069440		AI069440.1	GI:4949583					
Drosophila melanogaster (fruit fly)									
Drosophila melanogaster									
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.									
1 (bases 1 to 1101)									
Genoscope.									
Direct Submission									
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segret@genoscope.cns.fr									
- Web : www.genoscope.cns.fr)									
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).									
The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <a href="http://www.fruitfly.org">http://www.fruitfly.org</a> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Ooeegawa and Aaron Mammossier in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp. the same strain used for the library p1 and Est libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <a href="http://bacpac.med.buffalo.edu/drosophila_bac.htm">http://bacpac.med.buffalo.edu/drosophila_bac.htm</a> .									
location/Qualifiers									
1..1101									
organism="Drosophila melanogaster"									
/mol_type="genomic DNA"									
/db_xref="taxon:7227"									
/clone="BACR29P01"									
/clone_id="RPCI-98"									
/note="end : TE13"									
BASE COUNT	366 a	66 c	104 g	351 t	214 others				
ORIGIN									
Query Match		8.4%;	Score 52.4;	DB 29;	Length 1101;				
Best Local Similarity		35.5%;	Pred. NO. 3.5;						
Matches	151;	Conservative	65;	Mismatches	203;	Indels	6;	Gaps	1
QY	167	AAACATTTCTACATATATTTTACAGATCATTTGCGTTGAGGTTACAGACAGCTAG	226						
DB	552	AAAMWAATTTTTAAWMAATATAAAAAAATAATTTTAAATTTATATATTAATAAAGW	611						
QY	227	TATATGCTTATTTATATATGTTGCTCAAGATAGCTATTTATATGATGCCATATTC	286						
DB	612	AAAAAAWMTATTAHMTTAAATTTATATMTTAAWMTTAAATTAATTAATTTAAATTA	671						
QY	287	TTTTATATGCTGTACCGCAGTATTTCAATACGTCACATTTGGTGATATAGGACTA	346						
DB	672	TTTTTAAWMTTAAATTAATTAATAAAAAAATAATTTTAAATTTATATATTAATAA	731						
QY	347	CAAGGTTTTTAAATTTAAATGAGAAATTAATATATATAAAGGTTGAAGAAGATGT	406						
DB	732	ATPAATTTTTTTATTTAAWMTTAAATTAATAATTTAAATTAATAAATAAAAAATW	791						
QY	407	TTGCTTCTCAACA-----ACACTGTTCGACTTTACGATAGTTAGTTCAGATATGG	460						
DB	792	TATAAATTTAAATTAATTAATTAATAATTTAAATTAATAATTTAAATTTAAATTT	851						
QY	461	TTATAGGATATGTTCCATATATGATATATAAAGAGAGTTTCAACCTTTTACATTA	520						
DB	852	ATTTAAATTTTATTTTAAWMTTAAATTAATAATTTAAATTTAAATTTAAATTTAA	911						
QY	521	ATATGCAAGTCTGCTGAGGTGGCAGATTATATGCTTATACAGAAATAGTATCTGTT	580						
DB	912	ATTWGTATTTTATTTTAAATTTGATATATAAATAAATAATTAATGATATAAWMAA	971						

Qy	581	TATAT	585
Db	972	TGTAT	976
<hr/>			
RESULT 40	CNS00DKY	928 bp	DNA linear GSS 04-JUN-1999
CNS00DKY			
LOCUS			
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC # BACR27A24 of RPCT-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
ACCESSION	AL071865		
VERSION	AL071865.1 GI:4948170		
KEYWORDS	GSS.		
SOURCE	Drosophila melanogaster (fruit fly)		
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	Genoscope.		
AUTHORS	Direct Submission		
TITLE	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqrefgenoscope.cne.fr		
JOURNAL	Web : www.genoscope.cne.fr)		
COMMENT	Determination of this BAC-end and sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Oosagawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCT-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.		
FEATURES	Location/Qualifiers		
source	1..928		
	/organism="Drosophila melanogaster"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:7227"		
	/clone="BACR27A24"		
	/clone_lib="RPCT-98"		
	/note="end : 17"		
BASE COUNT	262 a	70 c	84 g 321 t 191 others
ORIGIN			
Query Match	8.4%	Score 52;	DB 29; Length 928;
Best Local Similarity	30.1%;	Pred. No. 4.2;	Mismatches 181; Indels 1; Gaps 1,
Matches 118; Conservative	92;		
Qy	219	ACAGTCAGTAAATGTTATTATTATTAATGTCACAAGAATGATTTATGATGCC	278
Db	490	AAATATAATAGTAAATTAATTTTTATKAWAAAAAAAATTTTTTTTWTWMAAA	549
Qy	279	AATATTAATCTTTATGCTGTGTTAGCGCANGATTTCAATGCTGACAAAGTTGGCAT	338
Db	550	AMTATTTTWTWTTTTTTTTTTTWAMTTTAAATTTATTTATTTAAATTAATTTATA	609
Qy	339	ATGACGCTCAAGGTTT-TAAATTTAAATGAGAAATTTAATTAATAAAGGTTGA	397
Db	610	WTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	669
Qy	398	AAGAAATGTTGCTTCTCAACAAACACTGTTGCATTTTCGTTAGTTAGTCAAGTAA	457
Db	670	AAAWAATTAATAATTTTAAATAATTAATAATAATAATAATAATAATAATAATAATA	729
Qy	458	TTGTTATAGTGATTTGTCATATATGATTAATAAAGGAGGATTTTCAAACCTTTTACCAT	517

Db 730 WATAMAAAATATWMMWMMWMAATWTDGNNNNNNMMWMMWMAAMWMAAAWMMWMMWMMW 789

Qy 518 TATATTATGCACTCCTTAGGTGGCGCATTATTCCTTAATACAGAACTAGTCTG 577

Db 790 AAAAAAAAAAAAAAAAAAAAAAADDDDDDDDDDWMAAKKKKKKKKKKKKKKKKKKK 849

Qy 578 TTTTATATAGCTAATTCCTATGACAATTA 609

Db 850 TKTTTGARWTTTTTTTTTTTTTTTTTTTT 881

Search completed: October 4, 2003, 05:13:17  
Job time : 1941 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2003, 18:26:46 ; Search time 2052 Seconds

(without alignments)  
12400.481 Million cell updates/sec

Title: US-10-010-160-1

Perfect score: 622

Sequence: 1 atgctgatgacccagctaa.....aattatgcagtcgagatc 622

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

```
Database :
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vi:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hggo_hum:*
40: em_hggo_mus:*
41: em_hggo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80.8	13.0	300150	1 AP004598	AP004598 Oceanobac
2	69.6	11.2	299511	1 AFGENOM	AE015941 Clostridi
3	68.8	11.1	16343	3 AFGENOM	U6178 Apis mellif
4	62.4	10.0	303249	1 AP001515	AE001515 Bacillus
5	61.6	9.9	1105	1 BSFLHB	X74121 B. subtilis
6	61.6	9.9	208780	1 BSUB0009	Z59112 Bacillus su
7	57.8	9.3	110000	2 PPMAL7PI_01	Continuation (2 of
8	56.2	9.0	94563	5 AC096885	AC096885 Danto rer
9	54.8	8.8	8056	6 AX598900	AX598900 Sequence
10	54.8	8.8	9052	6 AX345047	AX345047 Sequence
11	54.6	8.8	19237	3 AC116967	AC116967 Dictyoste
12	54.4	8.7	8056	6 AX599046	AX599046 Sequence
13	54.4	8.7	347050	3 PPA929351	AL929351 Plasmodu
14	54.2	8.7	310003	1 AE014016	AE014016 Buchnera
15	54.2	8.7	347050	3 PPA929351	AL929351 Plasmodu
16	53.8	8.6	257757	3 AE014837	AE014837 Plasmodu
17	53.6	8.6	1501	6 AX598988	AX598988 Sequence
18	53	8.5	945	6 BD092427	BD092427 Identific
19	53	8.5	10525	1 AE001502	AE001502 Helicobac
20	53	8.5	10860	1 AE000589	AE000589 Helicobac
21	52.8	8.5	7167	6 AX345303	AX345303 Sequence
22	52.8	8.5	11258	1 AE010546	AE010546 Fusobacte
23	52.8	8.5	106763	9 AP002091	AP002091 Homo sapi
24	52.6	8.5	254050	3 PPA929358	AL929358 Plasmodu
25	52.4	8.4	99003	2 AL390756	AL390756 Homo sapi
26	52.2	8.4	133112	2 BX323887	BX323887 Danto rer
27	52.2	8.4	341050	3 PPA929357	AL929357 Plasmodu
28	52.2	8.4	349980	6 AX344554	AX344554 Sequence
29	52	8.4	5297	6 AX251268	AX251268 Sequence
30	52	8.4	40324	6 AX458634	AX458634 Sequence
31	52	8.4	40324	6 AX458634	AX458634 Sequence
32	52	8.4	217879	2 BX005264	BX005264 Danto rer
33	51.8	8.3	5523	6 AX345066	AX345066 Sequence
34	51.8	8.3	15575	3 AY242996	AY242996 Antherea
35	51.8	8.3	73334	6 AX347026	AX347026 Sequence
36	51.8	8.3	73334	6 AX356493	AX356493 Sequence
37	51.8	8.3	110000	2 PPMAL7PI_12	Continuation (13 o
38	51.6	8.3	63513	2 AC074061	AC074061 Homo sapi
39	51.6	8.3	119555	9 BX004827	BX004827 Human DNA
40	51.4	8.3	10807	1 CPU88070	U68070 Chlamydomon
41	51.4	8.3	16351	1 AE001283	AE001283 Chlamydia
42	51.4	8.3	300512	1 AE016995	AE016995 Chlamydia
43	51.2	8.2	12025	6 AX346201	AX346201 Sequence
44	51.2	8.2	98299	2 AC133155_3	Continuation (4 of
45	51.2	8.2	156651	2 BX284638	BX284638 Danto rer

#### ALIGNMENTS

```
RESULT 1
AP004598 300150 bp DNA linear BCT 17-MAY-2003
LOCUS Oceanobacillus iheyensis HTE831
DEFINITION AP004598 BA000028
ACCESSION AP004598.1 GI:22777144
VERSION
KEYWORDS
SOURCE
ORGANISM Oceanobacillus iheyensis HTE831
Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
REFERENCE 1
Oceanobacillus iheyensis gen. nov., sp. nov., a deep-sea extremely
halotolerant and alkaliophilic species isolated from a depth of 1050
m on the Iheya Ridge
```

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
2  
Takami, H., Takaki, Y. and Uchiyama, I.  
Genome sequence of *Oceanobacillus iheyensis* isolated from the Iheya  
Ridge and its unexpected adaptive capabilities to extreme  
environments  
Nucleic Acids Res. 30 (18), 3927-3935 (2002)

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
3 (bases 1 to 300150)  
Takami, H., Takaki, Y. and Chee, G.  
Direct Submission  
Submitted (26-DEC-2001) Hideto Takami, Japan Marine Science and  
Technology Center, Deep-sea Microorganisms Research Group, 2-15  
Natsushima-cho, Yokosuka, Kanagawa 237-0061, Japan  
(E-mail: takami@jamstec.go.jp,  
URL: http://www.jamstec.go.jp/jamstec-bio/DEEPSTAR/Fresearch.html,  
Tel: 81-468-67-9643, Fax: 81-468-67-9645)

FEATURES  
source  
Location/Qualifiers  
1..300150  
/organism="Oceanobacillus iheyensis HTE831"  
/mol\_type="genomic DNA"  
/strain="HTE831"  
/db\_xref="taxon:221109"  
/note="halocolerant and alkaliphilic"  
28..1017  
/gene="OB1462"  
28..1017  
/note="OB1462"  
/gene="OB1462"  
/codon\_start=1  
/product="hypothetical conserved protein"  
/protein\_id="BAC13419.1"  
/db\_xref="GI:22777145"  
/translation="MGVMEHREYVLSKESIKGAIKPDGIYVDCITGGGSHLTIAS  
RLNNGQLFAPQDDKALAAADRLSTADRIYVQSNFRGLEBQKRGHIGVIGIL  
FDLVSSFDLDGDSFYNHDLDMRDSQHLAAEVNVEYERLIVLSPFSGYE  
EKFSQIARAKIAYREODIKITHLQVLEIKDAIPAPARKGHPAKRIFQALRIAVN  
DELVEFNNAHQARLIANVGRIVAVITFHSLEIDCKQAFKMSYDCKPTPLPIVPE  
SHQAFKLVTRKPIADTSELDENRRSRAKRLVIEKSEWHEFTYEGWKK"

gene  
CDS  
1017..1391  
/gene="OB1463"  
1017..1391  
/gene="OB1463"  
/note="CDS ID OB1463"  
/codon\_start=1  
/translation="hypothetical conserved protein"  
/protein\_id="BAC13419.1"  
/db\_xref="GI:22777146"  
/translation="MSVNHATKMETNTQYRPETQPVKKYKRTWITKGEKVIYAF  
AGVLMIAFVWVSYSSDTLRLRELQIQEGVNOQKIANEGLSPEKEELSPERIIN  
IAEENGKIQNTVEYKQVSFNN"

gene  
CDS  
1408..3648  
/gene="OB1464"  
1408..3648  
/gene="OB1464"  
/note="CDS ID OB1464"  
/codon\_start=1  
/translation="penicillin-binding protein 2B (cell-division  
septum)"  
/protein\_id="BAC13420.1"  
/db\_xref="GI:22777147"  
/translation="MKKRTTHMSSIIWIFVAFELITGSPWYIOANGSVQVQIN  
EMADQRTASTMSEKGIYDKMGWVLAIDPPTRMALIDVASENAEELHNSA  
ESTAKLAPLVLEEEETOEPIQNGLDNRRFVEGKQKQISQVQKEIDELNIPGI  
NFEERPMRYIPNGMFASHIIGFAREEKEETGEVHERITGVGMEDVWNDILGGKQY  
ISYQDLVNNKLLDPSSEVLIQLEPDGDIYLTIDQKITLLQEDVMTQADEKYNPTMNA  
IIMPKSGEIVAMSRPSVDPNPPTDVENWYNDPISTPEPESGTTWKIFTMAAIBEGV

YNGSEGFSGSYIQNEKIRAIQDHNNGEMGTITVDSEFSSSVASAKLIMKIDSE  
SYEYLQAFGEPEPTGIDLPGEAVGKISYNPDSKLRFAQSSVTPIQMKASAI  
VNGKMYKPVYVDKVDVSTGSDITSETOEYQGPISSETAEOKMKASVYNGKASAI  
GKPYQLEBYVWVGKGTGAEIPNPDGGGYLQRENVVFSFMGAAPIDDPQLMVAYSIQ  
PELEPDESGSTPLSFIKPNVVENSLHVMIDIPKXNEADIDQMEPDPVNVKSTSEVE  
EMTEFGNTITVVGSKVSSIRAGASVLPNEKVIITIDPPTMDIKGMLRMTL  
ANLIDQVESFGSGVTTQNDIVGNPLQSGDYLVLELGPSTDDSEBSSENEERT  
GDESE"

gene  
CDS  
3732..5663  
/gene="spovD"  
3732..5663  
/gene="spovD"  
/note="CDS ID OB1465"  
/codon\_start=1  
/translation="stage V sporulation protein D (sporulation  
specific penicillin-binding protein) (spore cortex)"  
/protein\_id="BAC13421.1"  
/db\_xref="GI:22777148"  
/translation="MKRVSTYITKRIYTVPLALIIAVIIGRLAYQVPLGDDLAG  
QANLWTRDITFAERGIYLDSEGEVAENVTAPVIVMPQITNPEQTAVQLAKILD  
VPQDRVHEITIKNTSSVWIRREGIKIDKQKALRELDLSGVYLAQSKRYYPNGDL  
SHVIGFTGIDNGMLGLEISYDKLKQNGSLISLSPAKGRLELDEVPYDGLN  
LQTTINTKVOTIIEBELDKATIKYNPQALAIANPKTGILGMSRPNPEYKQDV  
DASIRGNLPIWSTYVPEGSPFKITTLAALIEGAVNLEDRFHDGDIENGATTHCW  
KKGHDOSTYEVYQNSCNPGFVNLGRLGTEKIFSYIDAFRGKGTGIDQGGNGI  
LFOENVPVELATTSFGQGVSTPIQVAVAAVAVNGVLYKPYLADIEWIDPVNGET  
VEKEPQLOKRVISEDTSEIRVALSEVVAAGTROPVYVDGKVGKGTQKQKPNP  
GYMENNYISFIFGAPADDEIVVAVVANDKDTLOFGVVAAPIVGTIIDSIRAME  
VEPTTDGLEKESWPDQKVEVPLDIBSTQKLEPYLVDSVTVNGEDTIIIEGSPK  
GTMLSGSTVRIYLDENE"

gene  
CDS  
5802..7283  
/gene="murE"  
5802..7283  
/gene="murE"  
/EC\_number="6.3.2.13"  
/note="CDS ID OB1466"  
/codon\_start=1  
/translation="UDP-N-acetylmuramoylalanyl-D-glutamy]-2,  
6-diaminopimelate ligase (EC 6.3.2.13)"  
/protein\_id="BAC13422.1"  
/db\_xref="GI:22777149"  
/translation="MELKILLECTFFNNGEVCNCEITSLEMSRKITSSAVCIT  
GFTVDGHDVQAVKNGASAIETSKPLKEGVFIQVEDNRLAMLAAYVYPIK  
HFPLIGVINGKTTVTLDDKIPEYHQKAGVIGTIOVKIGESTFPIVNTPALBL  
QKTFVNRKDVKQKIMEVSHALDMGRVYGCVDIIVFTVLSQDHDVNDIDYLR  
AKSLIFPOLGNGYSEKERYAIINDDSSSHLGRSTAOHITYTSCKEATIMAKDIE  
LTASGIRFKLHSPGNTTIGSRLMGMNVNVMMAAASAKYPLNVIOALSIEG  
VNGRFPPIVGGANYSIVDPFAHPDPSLENYOTIKDPAKRNYYVAVCCGDRKKRP  
LMAVVALNYADHAVFTSDNRTEDPQALIDDMTAEIDANSSGVYVVDKREGIYAKIO  
SAQKODIVLNGKHEHYQIIIGTKYVDPDRVARNAIKQGE"

gene  
CDS  
7289..8656  
/gene="murF"  
7289..8656  
/gene="murF"  
/EC\_number="6.3.2.15"  
/note="CDS ID OB1467"  
/codon\_start=1  
/translation="UDP-N-acetylmuramoylalanyl-D-glutamy]-2,  
6-diaminopimelate-D-alanyl-D-alanyl ligase (EC 6.3.2.15)"  
/protein\_id="BAC13423.1"  
/db\_xref="GI:22777150"  
/translation="MLPTVOMLAEIPKHQKISTDTITIOVTRDTPSPSSNSLYPI  
VGERFDGHNFIPIDAIKGALIALMDKSIPLPESEWNSFPVFEVDITLMOQLANSYL  
RSVNPVTVGVTGSKGTTKQDLITSVYTKTGAAGGNLNHGIKPLTISMOPTPE  
VIVVEMSGPGEHELKSTABDPYVYTNIGESKISGKISIAKRLITIEGICS  
NGNFIIDDEBLKENTIDMDSOIIGIGTKESDINISVKKHPDPPTTEVYTPKPH  
VNLGHEHNAANPAFISIGALNINSHENIDLAENKLEVTSEHRELEAVGIVNCVHIIINDA  
YNAATGEMKASIQVKEMEGFONKVLGILBELGTYSYHHRFISYISRIEYIY  
YGEQAFYIYDEITKQHKPNVWHYIKQEDVVPISIKENHNNQTLILFASGGMKEPMFI  
KELQO"



```
gene
8709..9689
/gene="miray"
8709..9689
/gene="miray"
/EC_number="2.7.8.13"
/product="phospho-N-acetylmutamoyl-pentapeptide
transferase"
/transl_table=1
/codon_start=1
/protein_id="BACJ3424.1"
/db_xref="GI:22777151"
/translation="MDLTGLITIVAFELITVLSPIFPELRNFGSIREGPOS
HOKKTGPTMGMLIIFSIITSLIMASRTDEGINQVWMLIFVFGYGLGFDDFI
KVAKRNIGLTSKOKLFGQIILALFYLINQGSFVIYVPGELQDGMFYAVIV
IFMWGASNAVNITDGLAGTAAIAFGAFIIVANGIPDHVAIVAFSLVAGALG
FIVNAPAPKVPMDGSLAGALIAISILKELIILITIGSFVETLSVITQVLS
FKTTGKRVFKNSPLHHHVELGSEMRVITFWLGLFPAVLGYIEVGM"
9689..11038
/gene="murD"
9689..11038
/gene="murD"
/EC_number="6.3.2.9"
/locus_tag="CDS_ID_081469"
/codon_start=1
/transl_table=1
/product="UDP-N-acetylmuramoylalanine-D-glutamate ligase"
/protein_id="BACJ3425.1"
/db_xref="GI:22777152"
/translation="MNVLINPYQHVILVLGAKSGTAANYLONHIQVITINDMATL
EDATVQKLTQMGALVLSHEISVLDKEDLIVKNGIRYDVIVIEAQRIPIVISEV
ELVHILNOPYIGTGSNGKTTTTLITEMENSVKTVAGNIVAGTAEVASLQSD
EKWMELESFQLOGIDLOQFSTAVLNLFEHLDPHGSFENVYAKNIFESOKHXY
LIYVADDNWSAIIKTAETKVPSSSPPLADGAMDDFLYYDEKILAIRDIVLG
KHNNELIAATKAKNLGATNBEIVQVLTTFSSGKHLERPGVINGXYINDSATNI
LATKALAAFNKAVNLLAGLDRGNITBELPIYLAHVKAMVFESITGKLADGAVAN
IPVLEKALDVQHAVEVAFALTDQDTLLSPACASWQYKTFEERGDMFIOALRLX"
11116..12234
/gene="spovE"
11116..12234
/gene="spovE"
11116..12234
/gene="spovE"
CDS

Query Match 13.0%; Score 80.8; DB 1; Length 300150;
Best Local Similarity 48.3%; Pred. No. 7.7e-06;
Matches 226; Conservative 0; Mismatches 242; Indels 0; Gaps 0;

19 AAAACAGAGAAAGCAACCCGAAAGAGAGTCTGCTGAAAGGAGTGTCCCT 78
112665 AAGACAGAGAAAGCAACCCGAAAGAGAGTCTGCTGAAAGGAGTGTCCCT 112724
79 AATCAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 138
112725 AAAAGCCAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 112784
139 TATTCAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 198
112785 TTCGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 112844
199 TTTCGGTTGAGGTTACAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 258
112845 ATTATTCGGATCTTCACAGAAAGAGTCTGCTGCTGCTGCTGCTGCTGCT 112904
259 ATAGCTATTTTATGATGCCAATATTTATTTTATGCTGCTGCTGCTGCTGCT 318
112905 GCAAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 112964
319 CGTGTACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 378
112965 TTAGTCAAGTTAGGTTTATTTAGTCAAGGAGTCTGCTGCTGCTGCTGCT 113024
379 AATATATATTAAGGTTGAAAGAGTGTTCCTGCAACAACCTGTTCGACTTTTA 438
113025 GATCGATTCAGGAGTCAAGAGTATTTTTCATACGTCGATTAAGTGAATTTCTGAAA 113084
```

```
Qy 439 CGTAGTTAGTCAAGTAAATGTTATAGTATGTTCCATATMGATT 486
Db 113085 TCATGTTAAATAATGTTGTTTGGCAGATGTTACATTTATGATT 11312

RESULT 2
AE015941/C 299511 bp DNA linear BCT 06-FEB-2003
LOCUS Clostridium tetani E88, section 6 of 10 of the complete genome.
DEFINITION AE015941 AE015927
ACCESSION AE015941.1 GI:28203514
VERSION AE015941.1
KEYWORDS
SOURCE
ORGANISM
Clostridium tetani E88
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.
1 (bases 1 to 299511)
Bruggemann,H., Bauner,S., Fricke,W.F., Wierer,A., Liesegang,H.,
Decker,I., Herzberg,C., Martinez-Arias,R., Merkl,R., Henne,A. and
Gotteschalk,G.
The genome sequence of Clostridium tetani, the causative agent of
tetanus disease
Proc. Natl. Acad. Sci. U.S.A. 100 (3), 1316-1321 (2003)
2457253
PUBMED 12552129
REFERENCE 2 (bases 1 to 299511)
Bruggemann,H., Bauner,S., Fricke,W.F., Wierer,A., Liesegang,H.,
Decker,I., Herzberg,C., Martinez-Arias,R., Merkl,R., Henne,A. and
Gotteschalk,G.
Direct Submission
Submitted (31-OCT-2002) Goettingen Genomics Laboratory, Institute
of Microbiology and Genetics, Georg-August University,
Grisebachstr. 8, Goettingen 37077, Germany
Location/Qualifiers
1..299511
/organism="Clostridium tetani E88"
/mol_type="genomic DNA"
/strain="Massachusetts"
/sub_strain="E88"
/db_xref="taxon:212717"
complement (317..1060)
/locus_tag="CTC01397"
complement (317..1060)
/locus_tag="CTC01397"
/codon_start=1
/transl_table=1
/product="membrane spanning protein"
/protein_id="AA035956.1"
/db_xref="GI:28203515"
/translation="MYKEVWKKYIKGFLLPFLVQILFKSTGLNLAFTDYLFGKG
GSLIENSTFNOLIINDIIPISLEIFGALFYSPIIMTITSLNEXNISGEDEKESI
GPIRIYVLAITISAIISMVILFPGFTIILPFWIIMTYIRNIILPEBAHYVYNISY
GESIKKIELEKRYSGEIIIGIVTGLMVGWISGLNITTPKTNIAVIFISFIST
NIQVYFMFWSICREKAEKTELEBY"
complement (1312..2043)
/locus_tag="CTC01398"
complement (1312..2043)
/locus_tag="CTC01398"
/codon_start=1
/transl_table=1
/product="membrane spanning protein"
/protein_id="AA035957.1"
/db_xref="GI:28203516"
/translation="MYKEVWQKKKKKFLPFLVQILFEGVFRRTGDSNIDTYTL
IKETISWANSIGENAIANSITMLVFLVSSLFYFLVVISIVNORINKENF
EESIDYPTRYEGNIIIMTIIIVLVSGFIISPYVLISPIIVSMIDLKCGAY
LVYNSSIDELSDGVLDGQDYMRIIFLSFIITFIINLTNASTNINIIITYSMDLLK
VSIKFIYIMSMYICKEDVLRKK"
complement (2209..2793)
/locus_tag="CTC01399"
complement (2209..2793)
/locus_tag="CTC01399"
CDS
```



```

Db 269427 AAAAAAGTAAAGAGTTGGTTAGCATTTACT-TTGCTTAGTAGATCTTTAGTTAAGCTT 269369
QY 138 TTATTCAGGCGGTAATGGAGCGTCATTTTGAACAATTTTCTACTATATTTTACGAAATC 197
Db 269368 TTTTAGGGCAGATAGTACATAGAAAATTTACAAAATGTCATGACGTCCTTTTAAAGATTT 269309
QY 198 ATTTCGTTTGAAGTTACAGACAGTCAGTATATGCTTTATTTATTTATTTGTCACAGA 257
Db 269308 TTATGGGATGAGCTAAATATTAATAATTTATTTATATATCTTTAACCATGTTGATGA 269249
QY 258 GAT-AGCATTTTATGATGCGCAATTTACTTTTATGCTGTTACGCGATGATTCAT 316
Db 269248 GATCAGCATTTATTAATCTTAACTTAACTTAATAGATATATGCGATATTTGCA 269189
QY 317 TACGTCATAGTTGGTGCATTTATGACTACAAAGTTTAAATTTAATGAGATAT 376
Db 269188 ATTTTATACAAACAGATTCATTTTACAAAAGAACTTAAAGCAGATTTAAAAAT 269129
QY 377 TTAATATATATAAAGGTTGAAAGGATGTTGCTTCAACAAACCTTGTGACTTT 436
Db 269128 TAAATCCGTAAATGAGATTTAAAAAGATATTT---CTCCACAAGAACTTAGTGAGCTTG 269072
QY 437 TACGATGTTTGTTCATAGTATTTGTATAGGATTTGCTTCATATATGATTTAAAGAG 496
Db 269071 TAAAGATTTAGCCCTGTAAATGATAGGCGCTATGATAGATATAGATTTTAAAGAAA 269012
QY 497 AGTTTCAACCTTTTACCATTAATTAATGACAGCTTCAGTGGCAGATTAATATGTC 556
Db 269011 ACTATATAGAAATTTTAACTTAATATTTAAAGATTTCCATTAATGCTTAAAGCTTTT 268952
QY 557 TTAATACAGAAATAGTACTTGTTTATATACGCAATTCCTATGACATTAATTCAGTGC 616
Db 268951 AAGCATATGCTGTAAAGCATTTTATATATATACCTTGTATATGATTTTATAGCTGTG 268892
QY 617 CAGA 620
Db 268891 GAGA 268888

RESULT 3
AMFGENOM/ 16343 bp DNA circular INV 16-JAN-1998
LOCUS Apis mellifera ligustica complete mitochondrial genome.
DEFINITION L06178
ACCESSION L06178.1 GI:336279
VERSION 1
KEYWORDS ATPase; NADH dehydrogenase; complete genome; cytochrome b; cytochrome c oxidase; cytochrome oxidase; mitochondrial genome; transfer RNA.
SOURCE Apis mellifera ligustica (common honeybee)
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea; Apidae; Apis.
1 (bases 1 to 16343)
Crozier,R.H. and Crozier,Y.C.
The mitochondrial genome of the honeybee Apis mellifera: complete sequence and genome organization
Genetics 133 (1), 97-117 (1993)
JOURNAL MEDLINE
PUBMED 93114603
FEATURES
source
Location/Qualifiers
1..16343
/organism="Apis mellifera ligustica"
/organella="mitochondrion"
/mol_type="genomic DNA"
/sub_species="ligustica"
/db_xref="taxon:7469"
/sex="female"
/tissue_type="thorax"
/dev_stage="adult worker"
1..66
/product="tRNA-Glu"
/note="codon recognized: GAA"

tRNA
/anticodon=(pos:31..33,aa:Glu)
116..178
/product="tRNA-OTHER"
/note="codon recognized: AAU; anticodon unspecified"
221..286
/product="tRNA-Met"
/note="codon recognized: AUG"
/anticodon=(pos:252..254,aa:Met)
296..358
/product="tRNA-Gln"
/note="codon recognized: CAA"
/anticodon=(pos:326..328,aa:Gln)
360..429
/product="tRNA-Ala"
/note="codon recognized: GCA"
/anticodon=(pos:393..395,aa:Ala)
433..501
/product="tRNA-OTHER"
/note="codon recognized: AAU; anticodon unspecified"
502..1503
/codon_start=1
/translation="MFPNPKYHMFYPLITITPVLMMNSNFIOMWMEFGTILIS
LINIKSTNPKSLIYYSVYSISPLFPMIIVLSSISPTKDTPTNPMFLKIG
TFPHFMWYISYEMWKKQJFLSLIKPLPIYMWKMTIKNSITLPLTNSLYISF
YARFPTLKRLKLLACSTIFNSFYPIFILELNKMFIMIIIVSNYELLSELNKFNQ
NFMFMFKKQWNTPLFMENSMYPIFLSPVKMILFMWVSKAVNMLFPLMTSS
MEMIMVYIILKKEVFLKMFYKKNFIDDKDKMYHSYFALTLISNISPTLNFL"
/complement (1503..1571)
/product="tRNA-Cys"
/note="codon recognized: UGC"
/anticodon=(pos:1536..1538,aa:Cys)
/complement (1592..1659)
/product="tRNA-Tyr"
/note="codon recognized: UUU"
/anticodon=(pos:1621..1623,aa:Tyr)
1722..1793
/product="tRNA-OTHER"
/note="codon recognized: AAU; anticodon unspecified"
1794..3359
/codon_start=1
/translation="MMKMFSTHKNIGIYIITIALMSGMLGSSMGLIIRBELSPGS
WISNDQIVNTVYTSNLFMIFFMVMPELIGRQNTLIPMLGSPDMAFPMNNISFNL
IPRSLMLLSNLPYPSRGTCGTVPPLSATYIHSPPVDFATFSLMGSISIMSGL
NLNVTITMMKNSRBNYDQISLPWSVYITAILIWSLPVLAGITWLLDPRNNTSF
DPMGGDPLIYOHLEFWEFGHPEVYIILGFGHISHIVNBSGKEIPNLNMIYMLT
GIGELGFIWAHMFVGLVDVTRAYATATMIIVPTGIVKYSMLATYSGTLKNTL
SILMSGLPIMLEFTIGLIGLIMSNSSIDILIHOTYVVGHFHYVLSMGAVFAISSEI
HWPLITGLLINIKMLKIQIIMWPIGLVNTLFPFHQHLGMSMRPRYSVDVSYCNS
ISGMSMISINSMPILFIILBSISKRLLPFRNQSLEMLNPLPDLHSHLEPLL
IKNLNKSILIKF"
3355..3424
/product="tRNA-OTHER"
/note="codon recognized: AAU; anticodon unspecified"
3618..4295
/codon_start=1
/translation="MSTWPMFMEQSNSSVYADNLISFHNWMMIIMISTITVYIILD
LPMNKRSLNPLKNNHIEITWITPIIILIIICPSIKLIVLIDELVNPFSKISGH
QWTVSYEYEPENRNPEDSYMLNINNLNQRLETDNRVTPKIPURLITTSIDVHS
WVPSLGIKIVDAVGRINQNLNLSKRPGLFFQGCBSICGMNHSFPMIESTSFOYFL
NMVWKOI"
4294..4362

```

tRNA  
 CDS  
 /product="tRNA-Asp"  
 /note="codon recognized: GAC"  
 /anticodon=(pos:4325..4327,aa:Asp)  
 4370..4438  
 /product="tRNA-OTHER"  
 /note="codon recognized: AAU; anticodon unspecified"  
 4444..4602  
 /codon\_start=1  
 /transl\_table=5  
 /product="Atpase subunit 8"  
 /protein\_id="AAB96801.1"  
 /db\_xref="GI:829006"  
 /translation="MPQWMPKWFIFYIYLLIPLYFIMLINSMLIKTKINKETLKIK  
 LKKRWFW"  
 4584..5264  
 /codon\_start=1  
 /transl\_table=5  
 /product="Atpase subunit 6"  
 /protein\_id="AAB96802.1"  
 /db\_xref="GI:552442"  
 /translation="MKLILMNLFEMDPDPTSNLSNMLFMLEPLIIPSPIMWIOS  
 RIMFLMKILMFNMEFVKSKYQSNIIIFISLMYIMTINFSILPVFPLTSHL  
 LLMSLNTISGHFSLIYLIYNNYIMFLSHLPILNSPVLNMFNFWIIEILSIIRPWL  
 SIRSANLISGHILITLIGFISNFSITLPINLMIQMLLTLEFMSMQSYFISLL  
 ILYSESN"  
 5285..6064  
 /codon\_start=1  
 /transl\_table=5  
 /product="cytochrome oxidase subunit 3"  
 /protein\_id="AAB96803.1"  
 /db\_xref="GI:552443"  
 /translation="MKKNPPPHMTNSFWPIILISFSFNTLISIVVIYSSISGMFL  
 NINISILMMMLFRDILIRESTFOGHSMFTNLFKFSNITLISELNFISFWTFH  
 SINSINIBINMTKPKIKFPMPEIPLNPSIVSGFTVTLTSHYLLINMLKLS  
 YLLITLIGIYFTLQTLIEVNSPFCFSDSIYGSIFPMATGFHLAVLSIGIFLLSL  
 YRMNMIHFSNNMNNFELAIYMHFVDVIMELVTFYIILL"  
 6119..6184  
 /product="tRNA-Gly"  
 /note="codon recognized: GGA"  
 /anticodon=(pos:6149..6151,aa:Gly)  
 6185..6538  
 /codon\_start=1  
 /transl\_table=5  
 /product="NADH dehydrogenase subunit 3"  
 /protein\_id="AAB96804.1"  
 /db\_xref="GI:552444"  
 /translation="MKFIIFYPIFIIISILLINKPISLYKKKOYKSSPEFCGN  
 PITKNLPRLSPFLMTMWFLLIDVEIILLPIIFLYLKSSSTMVSIEMISFILLLIT  
 TILLENMNYLWLF"  
 complement(6572..6638)  
 /product="tRNA-Arg"  
 /note="codon recognized: AAG"  
 /anticodon=(pos:6602..6604,aa:Arg)  
 6734..6802  
 /product="tRNA-Asn"  
 /note="codon recognized: AAC"  
 /anticodon=(pos:6765..6767,aa:Asn)  
 complement(6810..6878)  
 /product="tRNA-OTHER"  
 /note="codon recognized: AUU; anticodon unspecified"  
 complement(6892..8556)  
 /codon\_start=1  
 /transl\_table=5  
 /product="NADH dehydrogenase subunit 5"  
 /protein\_id="AAB96805.1"  
 /db\_xref="GI:829007"  
 /translation="MIKMWVCGILLFESFLMLMSLYLLYLNKEFFPEKNITFYNM  
 KFPFLIILYKSKSIFLVSMTFSMIIYSISYMDISEKDRFLYMLILISMWML  
 IISPNMLISIGMDGLISYGLVIYMKMKSFSGMTYILNLRGDIIGLLINGMTIT  
 YVSNMLSEYKNNFEMWYIILMAFTRKAOIPEFTMLPMAMAPPVASIVHSSVLTVA  
 AGYILIRVYNLIDENYKATYMLASITMLPAGLVANPELDKTVAYASTISQCPMM  
 SMYLSITELVEPLHIFHAMFKSLMFCVGSYWHATYSIDIRMTYIGATYTFPKMSI  
 LIFSILISGFPFLVGYYSKDLIIENFPEKATYSIMNLIIIGITFTVSYSPRMILV

cRNA		TSKRLMNVYIKREKKIMCISMMNMISLIIYSKLIENLNINLGINLMITKXWVF KNIIVGLGNGFPIKILINNKIGFKRSEFLRNALIKIYYKKIIMMTTEVYIEKS IIEIISKPSVTLNLTIELKISNLMLMINITYLLIYLITYLLIPLINE" complement (8557, .8624) /product="tRNA-His" /note="codon recognized: CCA" /anticodon=(pos:8589..8591,aa:His) complement (8644,.9987) /transl_table=5 /codon_start=1 /product="NADH dehydrogenase subunit 4" /protein_id="AB96806.1" /db_xref="GI:552446"
CDS		/translation="NYLLILLMLNLMNSMYLPWLFPNNKMKNNLLIIGNLIINL LNLFLNMLNDWIYIFCNLSFNNTSYGLIMLIMITGLIPISLNNSLNCLPNNLLIM IYLLIFLNNLLLFLEFGGLLIPLYVKMGVSERNLSGFYLMFYTMELPML YIIYYLYIDYSLNFMLEMNLMLNFYILMSFLVKIPLYHGMLLKAVAP KYSMLIASIMEKLGGYGMFLMTIYKNEFILLOKLIVAINSGVLILSMICLSODM XIISLISIVHMGLMIMSMTPLKTIIGGYLMISIHGSSSLFLVNAVITYQTSR LMINKMGINPFMSBLLMFNLCSMSGSPPSINLISEYMLILIGMISMLKPMYLLM YCFSPSYSLYELPFIINHCKIPTMEKIKGILVEYEVLLHLMIPLMLPKLYFI" complement (9991,.10254)
CDS		
Query Match:	11.1%; Score 68.8; DB 3; Length 16343;	
Best Local Similarity	47.6%; Pred. No. 0.0027;	
Matches 202; Conservative	0; Mismatches 222; Indels 0; Gaps 0;	
Oy	171 AATTTCTACTAATTTTTACAGAAATCATTCGGTGTGAGGTTACAGACAGTCAGTATA 230	
Db	5755 AAGTTTAATTATTTATATATATTAATTAATGAATGCTTAATGTAATCCTGATGA 5696	
Oy	231 TCGTTTATTTATTTATGTCCTCAAGAGTAGCTATTTATGATGCCAATTTACTTT 290	
Db	5695 TACAAGAATTAAGAAATTTAGTAAAGGAATTTCTATGAGATTAATAAAAATTTAATATTTT 5636	
Oy	291 TATTGCTGTAACGCATGAGATTCATCAACGTGTAACAAGTGTGCATTATGACTACAA 350	
Db	5635 TGGTGTCATGTTATTAATTTTCAATATTTGGAGAAATTTGATGAAGAAAAATGTTCA 5576	
Oy	351 GGTTTTTAATTTAAATGAGATAATTTAATTAATAAAAGGTTGAAGGAATGTTTC 410	
Db	5575 AAAAATGAATAAAAAATTTAATCTGTAATAAATAAATAATTTCTCAAATTTTAA 5516	
Oy	411 TTCTCAACAAACCTGTTGCACTTTAACGATGTTTCCAGTATTTGTTAAGGAT 470	
Db	5515 AAAATTTGATATAATATTTGAATGTAAACCCTGGAATGTTCTTCTCGAAATAATATCTCG 5456	
Oy	471 TGTTCCATPATATGATTTAAAGAGAGATTTTCAACCTTTTACCATTATTTATGCAAG 530	
Db	5455 AAATCATATATTTTAATTTAAATTTGAATTTAATAATTTAATAATATTTGAGAT 5396	
Oy	531 TCCTTCAGGCTGCGAGATTATATGCTTAATATACAGAAATGACTGTTTATATACCT 590	
Db	5395 TGATCTTAATTAATTAATTAATTAATTTGAATGAGATTTATTAATCATTAATGATTAAT 5336	
Oy	591 AATT 594	
Db	5335 AATT 5332	
RESULT 4		
AF001515/c	303249 bp DNA linear BCT 01-FEB-2001	
LOCUS	AP001515	
DEFINITION	Bacillus halodurans genomic DNA, section 9/14.	
ACCESSION	AP001515 BA000004	
VERSION	AP001515.2 GI:12641879	
KEYWORDS		
SOURCE	Bacillus halodurans	
ORGANISM	Bacillus halodurans	
REFERENCE	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.	
AUTHORS	Takami,H., Nakasone,K., Hirama,C., Takaki,Y., Masui,N., Fuji,F., Nakamura,T. and Inoue,A.	

TITLE	An improved physical and genetic map of the genome of alkaliphilic	JOURNAL	Submitted (22-MAR-2000) Hideto Takami, Japan Marine Science and
JOURNAL	Bacillus sp. C-125	Technology Center, Deep-sea Microorganisms Research Group; 2-15	
MEDLINE	Extremophiles 3 (1), 21-28 (1999)	Natsushima, Yokosuka, Kanagawa 227-0061, Japan	
PUBMED	99184645	(E-mail: takami@jamstec.go.jp)	
REFERENCE	10086841	URL: <a href="http://www.jamstec.go.jp/jamstec-e/bio/DEEPSTAR/FResearch.html">http://www.jamstec.go.jp/jamstec-e/bio/DEEPSTAR/FResearch.html</a> ,	
AUTHORS	2 (sites)	Tel: 81-468-67-3695, Fax: 81-468-66-6364	
TITLE	Takami, H. and Horikoshi, K.	On Jan 31, 2001 this sequence version replaced gi:10174886.	
JOURNAL	Reidentification of facultatively alkaliphilic Bacillus sp. C-125	location/Qualifiers	
REFERENCE	to Bacillus halodurans	1. 303249	
AUTHORS	Biosci. Biotechnol. Biochem. 63, 943-945 (1999)	/organism="Bacillus halodurans"	
TITLE	3 (sites)	/mol_type="genomic DNA"	
JOURNAL	Takami, H., Nakasone, K., Ogasawara, N., Hirama, C., Nakamura, Y.,	/db_xref="taxon:8665"	
MEDLINE	Maui, N., Fuji, F., Takaki, Y., Inoue, A. and Horikoshi, K.	/note="alkaliphile"	
PUBMED	Sequencing of three lambda clones from the genome of alkaliphilic	104. .1102	
REFERENCE	Bacillus sp. strain C-125	/gene="BH2267"	
AUTHORS	Extremophiles 3 (1), 29-34 (1999)	104. .1102	
TITLE	4 (sites)	/gene="BH2267"	
JOURNAL	Takami, H., Takaki, Y., Nakasone, K., Hirama, C., Inoue, A. and	/codon_start=1	
MEDLINE	Horikoshi, K.	/product="penicillin-binding protein"	
PUBMED	Sequence analysis of a 32-kb region including the major ribosomal	/protein_id="BAB05986.1"	
REFERENCE	protein gene clusters from alkaliphilic Bacillus sp. strain C-125	/db_xref="GI:10174887"	
AUTHORS	Biosci. Biotechnol. Biochem. 63 (2), 452-455 (1999)	/translation="MKVHLITMEDWLSGIALMKKGEDIPFISLGFANPAERIPNE	
TITLE	5 (sites)	HHTRFGIASGCKLFTALICQLVEAGLSFPTPLSDWLDAPFNVTIHLHLTHTSGVP	
JOURNAL	Takami, H., Takaki, Y., Nakasone, K., Sakiyama, T., Maeno, G.,	DYFDEETDDEFEDLMKDVPMYHLRDLDFLPFGHAPMKPPGHRFHYNNLFTLGLG	
MEDLINE	Sasaki, R., Hirama, C., Fuji, F. and Maui, N.	VVESGCVTFQEVYEAANVFORAGHRESGSEYFADLPKATLPGYIDLEGGSKTILYL	
PUBMED	Genetic analysis of the chromosome of alkaliphilic Bacillus	PVIGSDGAGVVTAEEDMMKLMALREHLNETYTKLTPHVAHCEDDDYGYGVWVK	
REFERENCE	halodurans C-125	QODAIKRYHMGVDPVCFSASFYPSNGIVVVCNQGSAIVWMAIEALFESEA"	
AUTHORS	Extremophiles 3 (3), 227-233 (1999)	complement (1146. .1469)	
TITLE	6 (sites)	/gene="BH2268"	
JOURNAL	Takami, H.	complement (1146. .1469)	
MEDLINE	Genome analysis of facultatively alkaliphilic Bacillus halodurans	/gene="BH2268"	
PUBMED	C-125	complement (1146. .1469)	
REFERENCE	Extremophiles in deep-sea environments, 249-284 (1999)	/note="BH2268"	
AUTHORS	7 (sites)	unknown"	
TITLE	Takami, H., Maui, N., Nakasone, K. and Horikoshi, K.	/codon_start=1	
JOURNAL	Replication origin region of the chromosome of alkaliphilic	/transl_table=1	
MEDLINE	Bacillus halodurans C-125	/protein_id="BAB05987.1"	
PUBMED	Biosci. Biotechnol. Biochem. 63 (6), 1134-1137 (1999)	/db_xref="GI:10174888"	
REFERENCE	10427704	/translation="MKVITLGGSKFRKROEQAASAVLTIKGHVVLVSQFEPSGIEI	
AUTHORS	8 (sites)	TEBOVQFENHRKIMSDEIFPLVDVGYIGRSTRKEIEFATSKMPRIKYSSELA	
JOURNAL	Takami, H. and Horikoshi, K.	STTS"	
MEDLINE	Analysis of the genome of an alkaliphilic Bacillus strain from an	complement (1636. .2373)	
PUBMED	industrial point of view	/gene="BH2269"	
REFERENCE	Extremophiles 4 (2), 99-108 (2000)	complement (1636. .2373)	
AUTHORS	10805564	/gene="BH2269"	
TITLE	9 (sites)	/note="BH2269"	
JOURNAL	Nakasone, K., Maui, N., Takaki, Y., Sasaki, R., Maeno, G., Sakiyama, T.,	unknown"	
MEDLINE	Hirama, C., Fuji, F. and Takami, H.	/codon_start=1	
PUBMED	Characterization and comparative study of the rrm operons of	/transl_table=1	
REFERENCE	alkaliphilic Bacillus halodurans C-125	/protein_id="BAB05988.1"	
AUTHORS	Extremophiles 4 (4), 209-214 (2000)	/db_xref="GI:10174889"	
TITLE	10 (sites)	/translation="WVKKGLIVGVFALISGLMACNGOEBEVEKNIDGASGIPSI	
JOURNAL	Takami, H., Nakasone, K., Takaki, Y., Maeno, G., Sasaki, R., Maui, N.,	EEIIDIINNLPDPAATVNGESIIDEKTYVVALEQQMTAAOYGISLEGGKQMLISIEE	
MEDLINE	Fuji, F., Hirama, C., Nakamura, Y., Ogasawara, N., Kunata, S. and	QTIIEQLINELQQAATKEIEVESEBIEDELFAIVAFQSEBIEALMEALLEGSGSMBE	
PUBMED	Horikoshi, K.	LREIEIHYVQOQKVEEETEAIVNTEIEIIOARVEEKOYVTEDELPTPEELQSNIEQR	
REFERENCE	Complete genome sequence of the alkaliphilic bacterium Bacillus	LIAREQOELEVEPEKLRDGDIVVH"	
AUTHORS	halodurans and genomic sequence comparison with Bacillus subtilis	complement (2582. .3367)	
TITLE	Nucleic Acids Res. 28 (21), 4317-4331 (2000)	/gene="BH2270"	
JOURNAL	11 (bases 1 to 303249)	complement (2582. .3367)	
MEDLINE	Takami, H. and Takaki, Y.	/gene="BH2270"	
PUBMED	Direct Submission	complement (2582. .3367)	
REFERENCE		/note="BH2270"	
AUTHORS		/codon_start=1	
TITLE		/transl_table=1	
JOURNAL		/product="ABC transporter (ATP-binding protein)"	
MEDLINE		/protein_id="BAB05989.1"	
PUBMED		/db_xref="GI:10174890"	
REFERENCE		/translation="MADILQVEQTLTKRKKPILQIISWQEGEHNAIVLNGSGKT	
AUTHORS		SLIKITGVMPTEGAVVAGNKGHPIDEVKRIQSVMSLDRFHTKQDPLVLI	
TITLE		VLSGHTGVMPTEGAVVAGNKGHPIDEVKRIQSVMSLDRFHTKQDPLVLI	
JOURNAL		AKPEPLIDPEPTGDLIAREOLITLESVATTKGAPPLVYTHYPEIQLPITHVM	
MEDLINE		LKQGHVIAQGRKDEVLTKEMITTAFCGSLDVIAQSGRYWVKPM"	
PUBMED		complement (3378. .4310)	
REFERENCE		/gene="BH2271"	
AUTHORS		complement (3378. .4310)	
TITLE		CDS	

```

/gene="BH2271"
/codon_start=1
/transl_table=11
/product="dipeptidase"
/protein_id="BAB05990.1"
/db_xref="GI:10174891"
/translation="MRIFPAHCDALIKLWEDPARREYTDPAIEANAVRMEGVAAOC
FATFVSPNIPFDOKFOVALQOELHEKTLKHPQSFTERMDLTDADQIGALT
LEGADPFDDPTLKTLYLGVNSLGLTNQNALCADGEGPBGGLTEWGAVALN
NEHVLTVDVSHLSERGFMDVIEYAPFVATHSNKAICHRRLTDAQKALIEAGF
IGIVFHPHFLTGTDDASIDVMKGIHICSLGGQKVGSGDFDGI PCFVDGLEHAGR
FQWFMDELKRYDEELVRGPAWDFLKATPISTYV"
4406..4840
/gene="BH2272"
4406..4840
/gene="BH2272"
/codon_start=1
/transl_table=11
/product="heat shock protein class I (low molecular
weight)"
/protein_id="BAB05991.1"
/db_xref="GI:10174892"
/translation="MSRESGSKLPVNYODFLRSIDDFPSKSPQHLHNGMFAPSIP
IROYKEQPIIEALPVGSKNOIQLDVFPQYVRIHIEEVLMEDEDEKGIYSKIS
SOAREVIEPPYPVQEGDIKARLRNGLTLAIPIHNRKINIE"
4864..5193
/gene="BH2273"
4864..5193
/gene="BH2273"
/translation="unknown"
/codon_start=1
/transl_table=11
/protein_id="BAB05992.1"
/db_xref="GI:10174893"
/translation="MTNHLRQTVSGAYYQDPQHSHPNHPOPTSSGWMQMTSTG
PSEPPFYNYVFSQKPKPASPMKAFTNODGFINDIKTFQTVQVYKTVNQYTPLVKS
VGSIFIK"
complement(5202..5510)
/gene="BH2274"
complement(5202..5510)
/gene="BH2274"
/translation="note="BH2274"
Query Match 10.0%; Score 62.4; DB 1; Length 303249;
Best Local Similarity 49.0%; Pred. No. 0.017;
Matches 196; Conservative 0; Mismatches 201; Indels 3; Gaps 1;

```

```

DB 170831 ATAGATCCAAATTAAGAGAGCTTAACGAATTTCTCGTTC 170792
RESULT 5
BSFLHB 1105 bp DNA linear BCT 17-MAR-1994
LOCUS BSFLHB
DEFINITION B.subtilis flhb gene.
ACCESSION X74121
VERSION 1
KEYWORDS flagellar protein; flhb gene.
SOURCE Bacillus subtilis
ORGANISM Bacillus subtilis
REFERENCE 1
AUTHORS Carpenter, P.B., Zuberi, A.R. and Ordal, G.W.
TITLE Bacillus subtilis flagellar proteins Flp, FlpQ, FlpR and flhb are
related to Shigella flexneri virulence factors
JOURNAL Gene 137 (2), 243-245 (1993)
MEDLINE 94131291
PUBMED 8299954
REFERENCE 2 (bases 1 to 1105)
AUTHORS Carpenter, P.B.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-1993) P. B. Carpenter, University of Illinois, 190
Medical Sciences Building, 506 S Mathews, Urbana IL 61801, USA
FEATURES
Source Location/Qualifiers
1..1105
/organism="Bacillus subtilis"
/mol_type="genomic DNA"
/isolate="O11085"
/db_xref="taxon:1423"
1..1083
/gene="flhb"
1..1083
/gene="flhb"
/codon_start=1
/transl_table=11
/product="flagellar biosynthetic protein"
/protein_id="CA52218.1"
/db_xref="GI:395390"
/db_xref="SWISS-PROT:P35538"
/translation="MKLRVLDQFPAGEKTEKATPKRKRKDTKKGGQVAKSSDVNSAVL
LVIFLSIAIPVWRDLRSFIEFPEYSLMKSESNVHLTFSLKDMGMIAPIL
LVALVAGVSNVMOVGFLPSAIVLOPLKEDIPKGRIRYSMAIYELIKSIIKIV
VGRAPAVLMLHGEIIRLPILTEBALSFSKATLMGSLGAAULILADILYKQ
FDYENIKMSKQDIKDEYKSGDPIIKSKIKOROREWAMRMQEVKADVIITNPT
HYAIALKDYDEKMDAPYIVAGVDHLAKIRKIAKHEHVMVNRPLARALYDVEID
OAVDEEFKVALAEILAVYVTKOKQY"
BASE COUNT 355 a . 192 c 252 g 306 t
ORIGIN
Query Match 9.9%; Score 61.6; DB 1; Length 1105;
Best Local Similarity 47.4%; Pred. No. 0.12;
Matches 184; Conservative 0; Mismatches 204; Indels 0; Gaps 0;

```



```
/translation="MKKLRILTHIMYKLMKLGKSDYVYIGSEALPPPLSKDEE
OVLMLKPNQDAPARAILITERLRLVVTARKEFTGINTEDLSITGILKAVNT
NPEKIKLATYASRCIENEILMYLRNKKISEVSYPLINDWGNELLISDVIGTD
DDIITKDIENAVDDKLLKALKALQOLNEREKQIMEIRFLGVBEKTKQKXVADMGISOS
YISLRERKIIKRLRKEFNKMY"
6605..7387
/gene="sigc"
6605..7387
/function="late forespore-specific gene expression"
/alternate gene name: spoIIIG"
/codon start=1
/transl_table=11
/product="RNA polymerase sporulation-specific sigma factor
(sigma-G)"
/protein_id="CAB13407.1"
/db_xref="GI:2633906"
/db_xref="GOA:P19940"
/db_xref="SMIS-PROT:P19940"
/translation="MSRNKVEICGYDTSKLPYLKNEERKLPFQLODEGDSAREKLV
NGNKLVLAVTORFNNRGEVYDDLFQVCGIGLMKSIDNFDLSHNKFTVAVPMITGE
IRRYLRDNPNLRVSRSLDIAYKALQVERILSETSKPELADIARVLEPHEIVFA
LDAIODPVSLPEPIYNDGDPYVMDQISDERNTDSQWIEELALKEGRRLNDEKMI
LRKRFQCKTQMEVAEIEIGISOQVSRLEKAKAIKMNKHQ"
7535..8496
/gene="y1mA"
7535..8329
/gene="y1mA"
/function="unknown"
/alternate gene name: ABC transporter (ATP-binding protein)"
/codon start=1
/transl_table=11
/protein_id="CAB13408.1"
/db_xref="GI:2633907"
/db_xref="GOA:O31723"
/db_xref="SPTREMBL:O31723"
/translation="MILQDNVSLKRNKWLKIDHMKVEKENVLYGLNAGAKTAL
LNMCSYFPTSGEGRVLRGHEGKTELEKLRKIGVSAALQOKLPADSAFAPLAD
GAVASIGLYETPSKETREKALGLEEDLGAIRAPRYETLSQGEKQALIRALMAD
ELLILDEPVTCIDFLAREKLDITTYIANKENAPSLVTHHAERILPVPKALLKQ
GEVFGSGIKEMLTDQILSAFFDPPIHVMQDRPFLTRAPPTNA"
8472..8496
/gene="y1mA"
8531..9811
/gene="y1MB"
8531..9811
/gene="y1MB"
/function="unknown"
/alternate gene name: acetylornithine deacetylase"
/codon start=1
/transl_table=11
/protein_id="CAB13409.1"
/db_xref="GI:2633908"
/db_xref="GOA:O31724"
/db_xref="SPTREMBL:O31724"
/translation="WMDQIYSLQKVEEKEELIQLATLISYOTPARPANTESIOS
WLAGYINELGSIKMDYVPGDPAVNGVSKATGSDVSLIINGVLIIDQYGEBEK
HDPRIPIREKGLIRGASDMKGMACVLPAYKLRRESIELPGLIIQSYIGEVEG
AGTLECCRGYHADPAIVADTSDMHIQOQGVITGMIIKSSQTHDTRARNAFHAG
GTFGSALEKAKAIAGLGELEHNSIMKSYPGFPGTNTINPAIEGRARNAFIAD
CLMTTVAHPEKTHDQVALEIEDYVSDIDILRENRPVFKGSGSMEDGEIIF
PALEVDPGHPGVALTASHOKVKECPFIIDVSQSVTQDGMLYDAGIPCVIAPGLHN
AHSVNEKYSIOLVYEYTKIILDFIISWGRKEQ"
9894..10139
/gene="y1MC"
9894..10139
/gene="y1MC"
/function="unknown"
Query Match 9.9%; Score 61.6; DB 1; Length 208780;
Best Local Similarity 47.4%; Pred. No. 0.027;
Matches 184; Conservative 0; Mismatches 204; Indels 0; Gaps 0;
```

```
Oy 19 AAACAGAGAAAGCAACCCCGAAACGAGCTAGAGAGCTGTTCTGAAGGAGTGCCT 78
Db 107613 AAGACAGAAAAAGCCATCCGAAAAACGAGAGATACGCAAAAAAGGCGAGGTAGCC 107672
Oy 79 AAATCAGAGAGTTACTTAAGCATTAAGCTACTCAGACAGGAGTGTGGGCTGCTATT 138
Db 107673 AAAAGCTCGATGTCATACATCCGCTTCTTACTGGAATATTCCTTTCGCTATTGCA 107732
Oy 139 TATTACAGCGTAATGAGACGTCAATTTGAAACAAATTTCTACATATTTTACAGATCA 198
Db 107733 ATTGTCGATATGAGAGACAGGCTGCTGATTCATAGAAACATTTTATACGGAATCG 107792
Oy 139 TTTGCGTTGAGTTACAGCACTGATATACCTTATTTATTTATTTAGTTGCTCAAGAG 258
Db 107793 CTCACATGAAGCTCTCAGATCAATGATACATTTGTTGTCAGCTTATTTAAGGAT 107852
Oy 259 ATATCTATTTATGATGACCAATATTAATTTATGCTGTTTACGAGCATTTTCATTA 318
Db 107853 ATGGGCAATGATACCTCGACCGATCTTGCTGTTGCTGTTGAGAGAGTGTAAAGAAC 107912
Oy 319 CGTGTACAGTTGTCATTAATGACATCAAAAGGTTTAAATTTAAATGAGTAAATTT 378
Db 107913 TATATGAGAGTTGGGTTCTTCTGTTGCTGAGATGATACAGCTTGAAGAACTG 107972
Oy 379 AATATATTAAGAGGTTGAAGAGATGT 406
Db 107973 GATCCGATCAAGGCTTAAACGATAT 108000
```

RESULT 7  
PFMAL8P1\_01/c  
WPCOMMENT  
Sequence split into 14 fragments LOCUS PFMAL8P1 Accession AL844507

Fragment Name	Begin	End
PFMAL8P1_00	1	110000
PFMAL8P1_01	100001	210000
PFMAL8P1_02	200001	310000
PFMAL8P1_03	300001	410000
PFMAL8P1_04	400001	510000
PFMAL8P1_05	500001	610000
PFMAL8P1_06	600001	710000
PFMAL8P1_07	700001	810000
PFMAL8P1_08	800001	910000
PFMAL8P1_09	900001	1010000
PFMAL8P1_10	1000001	1110000
PFMAL8P1_11	1100001	1210000
PFMAL8P1_12	1200001	1310000
PFMAL8P1_13	1300001	1325595

Continuation (2 of 14) of PFMAL8P1 from base 100001 (AL844507 Plasmidium falciparum 3D7)

Query Match 9.3%; Score 57.8; DB 2; Length 110000;  
Best Local Similarity 45.8%; Pred. No. 0.16; 237; Indels 0; Gaps 0;  
Matches 200; Conservative 0; Mismatches 237; Indels 0; Gaps 0;

```
Oy 173 TTTTCTACTATATTTTACAGATCATTTGCGTTGAGGTTACAGACAGTCAGTATATG 232
Db 26171 TTGTTTTTTTGTGTTGTTGTTATGTTTTTTTTTTTCTTTGATATGTCATTTGAT 26112
Oy 233 CTTATATTTATTTATGTTGCTCAGAGATGCTATTTATTTAGATCCATATTAATCTTTTA 292
Db 26111 TTTGCTGTTGTTTCTCAGAGTATTTGATTAAGTATTTTGAATTTATATTTTGA 26052
Oy 293 TTGCTGTTAGCGCATGATTTTCATAGTGACAGTTGGCGATTTAGCTACCAAG 352
Db 26051 ATTGTTATCTTTTGTCTCAGCTATTTAAGTGTGTTTCTTATTAATTAATTTG 25992
Oy 353 TTTTAAATTTAAATGAGTAAATTTATATTAATAAAGGTTGAAAGGATGTTGCTT 412
Db 25991 TTTTATAGTTGATACCTATTTATTTGATTTATTTATTTATTTATTTATTTATTTAT 25932
Oy 413 CTCAACAACACTGTTGACCTTTTACGATTTAGTTCAAGTATTTAGTTATGATG 472
Db 25931 TGCCATAATTTATTTGTTGATATTTATTTATTTTAAATTAATTAATTAATTAATGAT 25872
```





/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 3711 a 371 c 371 g 3603 t

ORIGIN

Query Match 8.8%; Score 54.8; DB 6; Length 8056;  
Best Local Similarity 47.2%; Pred. No. 1.1;  
Matches 167; Conservative 0; Mismatches 187; Indels 0; Gaps 0;

QY 160 CATTGGAACATTTTCTACTATATTTTACAGATCATTTCCGTTGAGGTACAGCA 219  
DB 6351 CGTTAAATAATTAATTTATTAATTTATTAATTAATTAATTAATTAATTAATTAAT 6410  
QY 220 CAGTCAGTATATGCTTTATTTATTTATGCTCAAGATAGCTATTTATTTGATGCA 279  
DB 6411 ATATATATTTATTAATTTAATTTATTTATTTAATTTAATTTAATTTAATTTAAT 6470  
QY 280 ATATTACTTTTATTTGCTGTTACGGCAGGATTTTCATTAAGTGTACAGTTGTCATTA 339  
DB 6471 ATAAATAAATATCGTTTTTTTATTTTCGTTATTTAATTTTCGTAATAATTTTCGTA 6530  
QY 340 TGGACTACAAAGCTTTTAAATTTAATGAGTAATTTAATATATATAAAGGTTGAA 399  
DB 6531 AACGTTTATTTAATTTAATTAATTAATTTAATTTAATTTAATTTAATTTAATTTAAT 6590  
QY 400 GGAATGTTGCTTCAACAAACACTGTTTCAGCTTTTCAGTATTTAGTTCAAGTAAT 459  
DB 6591 AAATTTATTTATTTTATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 6650  
QY 460 GTTATAGGATTTGTCATATATGATTTAATAAAGAGCTTTTCAACTTTTAA 513  
DB 6651 AATTTTATTTATTTTATTTAATTTAATTTAATAAATAAATTTTATTTTACGTTTAA 6704

RESULT 10  
AX345047 9052 bp DNA linear PAT 01-FEB-2002  
LOCUS AX345047  
DEFINITION Sequence 118 from Patent WO200928.  
ACCESSION AX345047  
VERSION AX345047.1 GI:16492933  
KEYWORDS  
ORGANISM  
SOURCE  
synthetic construct  
synthetic construct  
artificial sequences.  
REFERENCE  
1 Olek, A., Piepenbrock, C. and Berlin, K.  
AUTHORS  
TITLE  
Diagnosis of diseases associated with the immune system  
JOURNAL  
Patent: WO 0200928-A 118 03-JAN-2002;  
EpiGenomics AG (DE)  
FEATURES  
source  
1. 9052  
Location/Qualifiers  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 2166 a 112 c 2104 g 4670 t

ORIGIN

Query Match 8.8%; Score 54.8; DB 6; Length 9052;  
Best Local Similarity 47.2%; Pred. No. 1.1;  
Matches 167; Conservative 0; Mismatches 187; Indels 0; Gaps 0;

QY 138 TTAATCAGGCGTAATGCAAGCATTTTGAACAATTTCTACTATTTTACAGAAATC 197  
DB 799 TTTTATGAGATGAAAAGAGTTAGTTTATTTAGTTTATTTATTTATTTGTTAT 858  
QY 198 ATTTCGTTTGAAGTTACAGACAGTCAGTATATGCTTTATTTATTTATTTGCTCACA 257  
DB 859 TTAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 918  
QY 258 GATAGCTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 317

DB 919 TTTGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 978  
QY 318 ACGTGTACAGTTGGTCGATTTAGCTACACAGCTTTTAAATTTAAATGAGTAATTT 377  
DB 979 ATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1038  
QY 378 TAAATATTAAGAAGGTTGAAGAAGCTTTCCTTCAACAACACTGTTGCACTTT 437  
DB 1039 TGTGAAATTAAGAGATTTTAATGAGTATTTATTTGTAAGTATTTTATTTTGAAG 1098  
QY 438 ACGTATTTAGTCAAGTATTTGTTATTTAGTATTTGTCATATATGATTTATTA 491  
DB 1099 ACGTATTTATTTTATTTTGAATTTTGTATTTAATTTTAAATTTTGTGATATGA 1152

RESULT 11  
AC116967 19237 bp DNA linear INV 12-MAR-2003  
LOCUS AC116967  
DEFINITION Dictyostelium discoideum chromosome 2 map 5481549-5500786 strain  
AC116967  
AC116967.2 GI:28828487  
VERSION  
KEYWORDS  
HTG.  
SOURCE  
Dictyostelium discoideum  
Dictyostelium discoideum  
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
REFERENCE  
1 (bases 1 to 19237)  
Gloekner, G., Eichinger, L., Szafanski, K., Pachebat, J., Dear, P.,  
Lehmann, R., Baumgart, C., Parra, G., April, J. F., Guigo, R., Kumpf, K.,  
Tunggal, B., Cox, E., Quail, M. A., Platzer, M., Rosenthal, A. and  
Noegel, A. A.  
Sequence and analysis of chromosome 2 of Dictyostelium discoideum  
JOURNAL  
Nature 418 (6893), 79-85 (2002)  
MEDLINE  
22092622  
PUBMED  
12097910  
REMARK  
The Dictyostelium Genome Sequencing Consortium  
2 (bases 1 to 19237)  
Baumgart, C.  
AUTHORS  
TITLE  
Direct Submission  
JOURNAL  
Submitted (04-APR-2002) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany  
3 (bases 1 to 19237)  
Baumgart, C.  
AUTHORS  
TITLE  
Direct Submission  
JOURNAL  
Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany  
4 (bases 1 to 19237)  
Baumgart, C.  
AUTHORS  
TITLE  
Direct Submission  
JOURNAL  
Submitted (12-MAR-2003) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany  
On Mar 4, 2003 this sequence version replaced gi:19920066.  
CDS predictions from Genepid do not necessarily reflect true genes.  
Further information is available from IMB Jena, Department of  
Genome Analysis  
(http://genome.imb-jena.de/dictyostelium/)  
and the University Cologne, Institute for Biochemistry I  
(http://www.uni-koeln.de/dictyostelium/project.shtml)

FEATURES  
source  
1. 19237  
Location/Qualifiers  
/organism="Dictyostelium discoideum"  
/mol\_type="genomic DNA"  
/strain="AX4"  
/db\_xref="taxon:44689"  
/chromosome="2"  
/map="5481549-5500786"  
complement(join(<160..641,723..790))  
/note="Genid exon scores (in order of location ranges):  
22..51, 0..50 - GSCJ\_ID dd\_03333"  
/codon\_start=1

CDS  
Agency : Deutsche Forschungsgemeinschaft (DFG).  
Location/Qualifiers  
1. 19237  
/organism="Dictyostelium discoideum"  
/mol\_type="genomic DNA"  
/strain="AX4"  
/db\_xref="taxon:44689"  
/chromosome="2"  
/map="5481549-5500786"  
complement(join(<160..641,723..790))  
/note="Genid exon scores (in order of location ranges):  
22..51, 0..50 - GSCJ\_ID dd\_03333"  
/codon\_start=1

CDs

product="hypothetical protein"  
/protein\_id="AA051103.1"  
/db\_xref="GI:28828494"  
/translation="MDLLKKTQINAILNHHFSNEIDTFYFGDIENLIVKLTIEEYF  
HGEFGKTVPPNHPHFLIDINDINKIKITISKLSNISNDNIYPLSKSDAYFINDVY  
KKTYESIEINLFWALENGALFAGYKLTINNDDNNNNKNNNGVNIIVDENEKKIKINSNNLVL  
CWNYIKSDGKSSNNLHTLPEFRCK"  
2131. .2736  
/note="GeneID exon scores (in order of location ranges):  
50.44 - GSCJ\_ID dd\_03273"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AA051097.1"  
/db\_xref="GI:28828488"  
/translation="MATIFDILNTLNNNFNNKRCQCPINKSNKTIIDIIIPSMVMT  
NDKLIETRELASIKSDDIHDIKDSILITIOGKKKNLKKQOQVSEKSSISPTLSS  
KEDPSIEFEDDIKPKASTITTTITKENKODEKTKTSORKTISERSFNNFRKYIDL  
LTKVLYQDLNLSINQFENGGLTTITINKLHVSNTIKININ"  
complement(join(12919..3600,3675..4045))  
/note="GeneID exon scores (in order of location ranges):  
29.77, 4.69 - GSCJ\_ID dd\_01077"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AA051098.1"  
/db\_xref="GI:28828489"  
/translation="MLKENKKSLIHLIKDKSPFKTKGPPQDFPIENOIKIGLYEF  
FELDKNDSPVVISILNTNFIKFLPTLNENIDFELIVKVISIDIMKGYFNSFAIEM  
BDLKTNNIDNALNHHFSNEINTFYPDIDNLPILVPIKEYPQEGEYVKTPEPNHY  
IIDNNINKIKITISKLSNISNDNIYPLSKSDAYIDNRYKKTESINELMLAE  
NGLAGYKLTINNDDNNNNKNNKNNKNNKNTIYFDEDEREKYKNSDLVCMNFKGSDG  
KSSNLTYPFERCKGVALKVAKLTLSIENNIIPLCSTHDNINSILIEKGFKYS  
YSMAVVISKNNKINFE"  
join(5458..5626,5684..6099)  
/note="GeneID exon scores (in order of location ranges):  
4.76, 32.09 - GSCJ\_ID dd\_01076"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AA051099.1"  
/db\_xref="GI:28828490"  
/translation="MATIFDILNTLNNNFNNENCKRCOSTNKSNTIIDLIPSMVDYTM  
TNDKLIETRELAGEKKNLKKQOQOQKOLVIEKTSSTTSLDRKEDASIEEPD  
TDIKPSKSDNTPATTTTENEKDEKTKSTDPKPISERSFNNFRKYIDLTKLYQ  
LDLNSINQFENGGLTTITINKLHVSNTIKININ"  
complement(join(6342..6990,7089..7459))  
/note="GeneID exon scores (in order of location ranges):  
31.71, 5.21 - GSCJ\_ID dd\_01075"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AA051100.1"  
/db\_xref="GI:28828491"  
/translation="MLKENKKSLIPIKDKSFFKTKGPPQDFPIENOIKIGLYEY  
FTDNDSPVVISILNTNFIKFLPTLNENIDFELIVKVISIDIMKGYNSINALEM  
BDLKTNNIDNALNHHFSNEINTFYPDIDNLPILVPIKEYPQEGEYVKTPEPNHY  
IIDINDINKIKITISKLSNISNDNIYPLSKSDAYIDNRYKKTESINELMLAE  
NGLAGYKLTINNDDNNNNKNNRVVDENNNYVNNNLVCMNFKSGKSSNLTPIPE  
FRCKGIALKVAKLTLSIENNIIPLCSTHADNTKNSLIEKGFKYSYIAVWVISK  
NKNIN"  
join(8484..8774,8822..9133)  
/note="GeneID exon scores (in order of location ranges):  
9.50, 12.11 - GSCJ\_ID dd\_01074"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AA051101.1"  
/db\_xref="GI:28828492"  
/translation="MATIFDILNTLNNNFNNENCKRCOSTNKSNTIIDLIPMDVTT  
NDKLIETRELAGISKQDIHDIKDSILITIOGKKNLKKQOQVSEKSSISPTLSS  
LKPPEDDIKPKSDNTRVSTTTATKENKDEKTKSSDPKPISERSFNNFRKYIDL  
TKVLYQDLNLSINQFENGGLTTITINKLHVSNTIKININ"  
complement(join(9436..11722,11853..12590,12667..12816,  
12894..12974,13087..13307))  
/note="GeneID exon scores (in order of location ranges):  
170.09, 23.78, 2.31, 7.81, 7.70 - GSCJ\_ID dd\_01073"  
/codon\_start=1

CDs

product="hypothetical protein"  
/protein\_id="AA051103.1"  
/db\_xref="GI:28828494"  
/translation="MDLLKKTQINAILNHHFSNEIDTFYFGDIENLIVKLTIEEYF  
HGEFGKTVPPNHPHFLIDINDINKIKITISKLSNISNDNIYPLSKSDAYFINDVY  
KKTYESIEINLFWALENGALFAGYKLTINNDDNNNNKNNNGVNIIVDENEKKIKINSNNLVL  
CWNYIKSDGKSSNNLHTLPEFRCK"  
2131. .2736  
/note="GeneID exon scores (in order of location ranges):  
50.44 - GSCJ\_ID dd\_03273"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AA051097.1"  
/db\_xref="GI:28828488"  
/translation="MATIFDILNTLNNNFNNKRCQCPINKSNKTIIDIIIPSMVMT  
NDKLIETRELASIKSDDIHDIKDSILITIOGKKKNLKKQOQVSEKSSISPTLSS  
KEDPSIEFEDDIKPKASTITTTITKENKODEKTKTSORKTISERSFNNFRKYIDL  
LTKVLYQDLNLSINQFENGGLTTITINKLHVSNTIKININ"  
complement(join(12919..3600,3675..4045))  
/note="GeneID exon scores (in order of location ranges):  
29.77, 4.69 - GSCJ\_ID dd\_01077"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AA051098.1"  
/db\_xref="GI:28828489"  
/translation="MLKENKKSLIHLIKDKSPPFKTKGPPQDFPIENOIKIGLYEF  
FELDKNDSPVVISILNTNFIKFLPTLNENIDFELIVKVISIDIMKGYFNSFAIEM  
BDLKTNNIDNALNHHFSNEINTFYPDIDNLPILVPIKEYPQEGEYVKTPEPNHY  
IIDNNINKIKITISKLSNISNDNIYPLSKSDAYIDNRYKKTESINELMLAE  
NGLAGYKLTINNDDNNNNKNNKNNKNNKNTIYFDEDEREKYKNSDLVCMNFKGSDG  
KSSNLTYPFERCKGVALKVAKLTLSIENNIIPLCSTHDNINSILIEKGFKYS  
YSMAVVISKNNKINFE"  
join(5458..5626,5684..6099)  
/note="GeneID exon scores (in order of location ranges):  
4.76, 32.09 - GSCJ\_ID dd\_01076"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AA051099.1"  
/db\_xref="GI:28828490"  
/translation="MATIFDILNTLNNNFNNENCKRCOSTNKSNTIIDLIPSMVDYTM  
TNDKLIETRELAGEKKNLKKQOQOQKOLVIEKTSSTTSLDRKEDASIEEPD  
TDIKPSKSDNTPATTTTENEKDEKTKSTDPKPISERSFNNFRKYIDLTKLYQ  
LDLNSINQFENGGLTTITINKLHVSNTIKININ"  
complement(join(6342..6990,7089..7459))  
/note="GeneID exon scores (in order of location ranges):  
31.71, 5.21 - GSCJ\_ID dd\_01075"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AA051100.1"  
/db\_xref="GI:28828491"  
/translation="MLKENKKSLIPIKDKSFFKTKGPPQDFPIENOIKIGLYEY  
FTDNDSPVVISILNTNFIKFLPTLNENIDFELIVKVISIDIMKGYNSINALEM  
BDLKTNNIDNALNHHFSNEINTFYPDIDNLPILVPIKEYPQEGEYVKTPEPNHY  
IIDINDINKIKITISKLSNISNDNIYPLSKSDAYIDNRYKKTESINELMLAE  
NGLAGYKLTINNDDNNNNKNNRVVDENNNYVNNNLVCMNFKSGKSSNLTPIPE  
FRCKGIALKVAKLTLSIENNIIPLCSTHADNTKNSLIEKGFKYSYIAVWVISK  
NKNIN"  
join(8484..8774,8822..9133)  
/note="GeneID exon scores (in order of location ranges):  
9.50, 12.11 - GSCJ\_ID dd\_01074"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AA051101.1"  
/db\_xref="GI:28828492"  
/translation="MATIFDILNTLNNNFNNENCKRCOSTNKSNTIIDLIPMDVTT  
NDKLIETRELAGISKQDIHDIKDSILITIOGKKNLKKQOQVSEKSSISPTLSS  
LKPPEDDIKPKSDNTRVSTTTATKENKDEKTKSSDPKPISERSFNNFRKYIDL  
TKVLYQDLNLSINQFENGGLTTITINKLHVSNTIKININ"  
complement(join(9436..11722,11853..12590,12667..12816,  
12894..12974,13087..13307))  
/note="GeneID exon scores (in order of location ranges):  
170.09, 23.78, 2.31, 7.81, 7.70 - GSCJ\_ID dd\_01073"  
/codon\_start=1

```

/product="hypothetical protein"
/protein_id="AA051102.1"
/db_xref="GI:28828493"
/translation="MISLKEFLPLFLNFIQLSVNPISPEYNCLSTLDYDNLK
FKPDNNQYDICTPAMNDIVCTTNNINIKLANSQTDNIIDLPLNVCNCPQWDELI
VNLKATPNIYYKPEANTVRYSNIGLIGTLPAPRALIEFSDINFTFMQ
YLDKATYSLIMSSQSYVYLNDGNSGSSILTNRIISNLSIDLSYKRIQVTRK
ISPRDYSTISFLKLIQITDQSGMLTITTYRDNATLDQITPNSITPIYVSAIFPK
PEFSDITLFLKNIYKISITLVPETSGNSFPLIYKNIIEISFIDGYLTPNNLN
NDDITTLNPNLSIKSEQPEPPLSGSDPDTGTGTFGLTIDKSCWTEIAVSLNLTGL
VPSCAQENNTVNSYSKLTQKSMFRRASNFNSNTPTCTTFEPKELIDGIEFL
NKWATGNDIGDVTMLKINTLISLDEPTIYVINGNGYATVAGANNLKXVEFVY
LFTTTPNNVYATPEPTNNNIPIIKSVLISKSTNGEIOVSGHFSYLGPTNOVOIANG
NOSTTSSDPFNAOCTMESGKOLPIISNTIPILOIISIKDPSYKATIKESNFDDN
VCPNDCTSVGNSTCDLSSGTCGLPFTIGSCLIGESVDPQSGVGNCTTIGECTI
CDSSRGSDSCILFIEDPLDCNNSNGVCTPTKCECKENMSGAPTLIYHYITSTI
PSTNGVASPILFEPGDTNNLISITIGNLPCPLIYSLVNTCAITAPGPFVLLVVVO
NNIITSYDQYVLTANNIQTCPNPKCSNGQCTNLTNGCQCNNGENADCSGIIINGGN
NGNSDSDGNGNNGPDTGVDGPGTNGNTISNOEVOFQIFKSLVEIDENGNVATY
SLRNMWTNSNKSNEIYTLISOTLOSNIISIKIEIENMDEKXETFADETPNLAOG
SIKESGIYVNSYKNNIANTLKIELVSDIOINSDPMECHMPSKNTINIKNESTN
YIKISKNKILIEGRITNKLVSDGRPTIYLTSMKNEKDSISTIEDLPHCIESCTIDP
FVSLISPEFKNECNSKDGKRSYVLPVAVASVVSAGALGFLIYRKKGESVLLNK
LKIKNK"
complement(join(13861..17107,17185..17904,18265..18346))
note="Geneid exon scores (in order of location ranges):
246_52, 30_89, 0_31 - GSCI_ID dd_01072"
/codon_start=3
/product="similar to Podocoryne carnea. EGF-like protein
(fragment)"
/protein_id="AA051104.1"
/db_xref="GI:28828495"
/translation="WFLIPGCTNNILPESNFTCEPTLVAELTHISNDINCTLRLEH
ISGIVNLIDISPNVYVYVNLNSDGNBSCTVLEPNVANNIYSPSKYKNNILQEST
IFPSTINPMKTISSNQSLSTRALQINDIYLDNKAQVSRFYASLITADIPETVDMSSF
VNHVITATLFPNGLYHLNGALPIIPESCDNFELSGNLVFPNNLNMGYSVTSFEN
SIKNTQFPTEKLGSGSYVFTNPSNLVFTVGNWSCTTELISVETGMLTGILIPCAFQY
FNVTIINSYSKLTQKSMLTRPFSKNSFNNTPTCTTFEPKELIDGIEFLKNTVND
DIGEPTVTLKINTLISLDEPTIYVINGNGYATVAGANNLKXVEFVLPFTPNV
TYTPTPNINIPILIKSVLISKSTNGEIOVSGHFSYLGPTNOVOIANGVCSITINS
DPEGACCTMESGKOLPIISNTIPILOIISIKDPSYKATIKESNFDDNVCNDCT
SVGNSTCDLSSGACKLPGFVSDGLIEGVSDQSGHGDGLIIECTICNSYQCS
DCLPLIPCPIYGSIPCNNGSGWTCNNQGTICYDSSHONGDGLPIIECPVSGLPQN
GGLANTCNKTKTQCGDSHONGDGLPIIECTIVSGLPNCGSLANTCNMQGTQCGDS
HOGSDCSPLIECPVSGLPNCGSLANTCNMQGTQCGDSFPGNSCSLPIVCPPIFES
LPNCGSLANTCNMQGTQCGDQGNOSDGLPIVDCPTNGLPNCGSLANTCNMQGTGN
CDSHOGSDCSPLVDCPESNGLPCNGLANTCNMQGTICNCSHONGSVPEVER
MTTGVACNGSLANTCNMQGTQCGDQGNOSDGLPIVDCPDNDNSNGVCDTIRKGCN
CKENWSGPTCLIPHYITSTIPSTNGVAVLIGFPGTHNNLSITIGLPCPLLN
STYVNTAPGPFVLLVVVOANNITVYSNKOYALINQTCENKCSNGQCTNLTNGC
CKNNNGFADSDGIIINGGNNGNSNSDSDGNNEPDTGVDGTGTTISIKSE
OQLEPFKSLVEIDENGNVATYSLNSMTSDTNEIYTLISOTLOSNIISIKIE
EINMDEKXETFADETPNLAOGSIKESGIYVNSYKNNIANTLKIELVSDIOINSDN
BEINMDEKXETFADETPNLAOGSIKESGIYVNSYKNNIANTLKIELVSDIOINSDN
KDSISISDLPHCIESCTIDPFSVLSIPEFKNECNSKDGKRSYVLPVAVASVVG
SALIGLFLIYRKKGESVLLNKIKTKNK"
BASE COUNT      7622 a      2154 c      2055 g      7406 t
ORIGIN
Query Match      8.8%  Score 54.6; DB 3; Length 19237;
Best Local Similarity 47.1%; Pred. No. 0.98;
Matches 168; Conservative 0; Mismatches 189; Indels 0; Gaps 0;
172 ATTTTCTACTATATTTTACAGAAATCATTCGATTGAGATTACAGACAGTCAGATAT 231
Db ATTTTCTACTATATTTTCTTATAACATATTTATTTATTTATTTATTTATTTATTTAT 4077
172 ATTTTCTACTATATTTTACAGAAATCATTCGATTGAGATTACAGACAGTCAGATAT 231
Db ATTTTCTACTATATTTTCTTATAACATATTTATTTATTTATTTATTTATTTATTTAT 4077
232 GCTTATTTATTTATGTCCTCAAGATAGCTATTTTATGATGCCAATATTACTTTT 291
Db TATTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 4137
292 ATTGCTGTACGGCATGATTCATTTAGCGTGTACAGAGTTGTCATATGACATCAAG 351

```







CDS  
 /locus\_tag="bhp006"  
 4942..6486  
 /gene="atpA"  
 /locus\_tag="bhp006"  
 /EC\_number="3.6.3.14"  
 /note="COG0056; FOF1-type ATP synthase alpha subunit"  
 /codon\_start=1  
 /transl\_table=11  
 /product="ATP synthase subunit alpha"  
 /protein\_id="AAO26750.1"  
 /db\_xref="GI:27903916"  
 /translation="MQISSESEICELISKIKKPIISSMNEGVISVSDGIITVGL  
 SDVWQSELSLPDQKVAIALNLKRVGALVWGEYTHITTEGKLIISTGRIFEIVSGK  
 FLGRVINALGPIDGKRIQDEKFLPVEINAPGVIDROKISEPLQGTGKSIDANVPIG  
 KGRRELIIGDROGKTSIAIDTIINORTKIKCIYVIAIGKSTIVNLVROLDNOAL  
 NHTIVIVASASEAALQYLVYSGSGSGEPFDGKALIVYDDLKHAIVROISLL  
 LRPRGEAPRPGDIPEYHARLLEBACRVNSYLNKLTGKTVKSTGSLTALPIETED  
 GDVSSFTPTNVISTIDGQIFLESNLVNSGIPALNPGISVRVGAACQCIKRLSG  
 IRTLAQYQELIARQFSESLDEITRNQLHGKLIILKQKQYHPMSIAQAILFA  
 AENNRLDVSEVQIIEKEMLLFENTNSELVLSINNYKINDVVENQLSDVINAEK  
 LTKCMS"  
 6524..7399  
 /gene="atpC"  
 /locus\_tag="bhp007"  
 6524..7399  
 /gene="atpC"  
 /locus\_tag="bhp007"  
 /EC\_number="3.6.3.14"  
 /note="COG0224"  
 /codon\_start=1  
 /transl\_table=11  
 /product="ATP synthase gamma chain"  
 /protein\_id="AAO26751.1"  
 /db\_xref="GI:27903917"  
 /translation="MIGREIRSKCKSINNTKITKAMEMVSIKAKIKRMCSSRP  
 YPTINQVISHVITGNLEHHTYFNQKVKIGVIVSTDGKLCGNLTIDLFKVLVY  
 LTERINHIILNLFVIGTKALTFKSPFNINVFELSKNDPKIIDMEMIRISLEMY  
 ISGKIDKFLAVNKNSTIIQTPILVOLLLPKRLGKKEVKTWDYIYESNSKVLN  
 VLNRYEFOIYOSILENLVCEQASRLAMKQATDNSADLLKALQNNVKNYRQSSIIQ  
 ELTEIISGAAVSLN"  
 7427..8824  
 /gene="atpD"  
 /locus\_tag="bhp008"  
 7427..8824  
 /gene="atpD"  
 /locus\_tag="bhp008"  
 /EC\_number="3.6.3.14"  
 /note="COG0055"  
 /codon\_start=1  
 /transl\_table=11  
 /product="ATP synthase beta chain"  
 /protein\_id="AAO26752.1"  
 /db\_xref="GI:27903918"  
 /translation="MITGKIVQIIGAVVDFEQQSVPKIFNALKVDNQSILILEVQ  
 QQLSGSIVRTIAMSSNGKRLVLDLEHGIKVPVGTATLGRIVNVLGQIDMKGFL  
 KMHNSDIETVIRKAPSYSEQLTSEYLETGIKVIDLIPFGKGVGFLYGGAGG  
 KTVMMELIRIATIEHSGSYFTGVEGTRGNDPYHMSDRVLDRKSLVYQGNMP  
 PGNRLVAFGLRTIAEKPRNGHDVLLPTINIVRYTLAGTYSALLGRISAVQORT  
 LSEMGVLOERITTSKGSITSIAVYVPADLDLDPSPATTFSHLDSTITLSROVSL  
 GIYPAIDPLNSTSQDPRIVGQHYVAVGVRSLIRYQELKQIILILGDELSESD  
 KILVSRARKIQKFLSQPFVAIEFTFGSKYVKLQDTINFKDIIIEKQVHVPQARY  
 MWGSEINVEIKSKL"  
 8826..9263  
 /gene="atpC"  
 /locus\_tag="bhp009"  
 8826..9263  
 /gene="atpC"  
 /locus\_tag="bhp009"  
 /EC\_number="3.6.3.14"  
 /note="COG0355"  
 /codon\_start=1  
 /transl\_table=11

Query Match 8.7%; Score 54.2; DB 1; Length 310003;  
 Best Local Similarity 46.4%; Pred. No. 0.53;  
 Matches 213; Conservative 0; Mismatches 243; Indels 3; Gaps 1;

Oy 163 TTGAAACATTTTCTACIATATTTTACAGAACATCTTGGTTGAGTTACGACACG 222  
 |||||  
 Db 261562 TTTAAAGAAATTTTATTTTAAAGTTTACCTTTTAAATCTTATTTTCAAAAGATTTA 261621  
 |||||  
 Oy 223 TCAGTATATGCTTATTTATTTATTTATGTCGCAAGAGATAGCTATTTATGATGCAATA 282  
 |||||  
 Db 261622 TTTTATGATCACAATTTCTTATATGATTTTTCACAAATATCTTATTTCTATTTAGGATA 261681  
 |||||  
 Oy 283 TTACTTTTATGCTGTTACGCGATGATTTTCAATGATGATGATGATGATGATGATGATGATG 342  
 |||||  
 Db 261682 ATTTTTCCTATGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 261741  
 |||||  
 Oy 343 ACTCAAAAGTTTAAATTTAAATGAGATTAATTTATTTATTTATTTAAAGGTTGAAGGA 402  
 |||||  
 Db 261742 AATTTTAAATTTTAAATTTGATATTTAAATTTAAATTTAAATTTAAATTTAAAGAT 261801  
 |||||  
 Oy 403 ATGTTGCTTCTCAACAAACCTTGTGACTTTTACGTAGTTAGTTCAAGTATGTT 462  
 |||||  
 Db 261802 GTGTTTCTATTAAGATATTTGATTTATTTTAAACAGCTTTAA--AATAGTTTT 261858  
 |||||  
 Oy 463 ATAGATATGTTCCATATATGATTTAAAGAGAGTTTCAACCTTTTACATATATAT 522  
 |||||  
 Db 261859 GTGATATTTAGATATTTATTTATTTATTTATTTATTTAAAGATATGTTTCTCA 261918  
 |||||  
 Oy 523 TATGCAAGTCTTACAGTGTGCGAGATTTATGCTTAAATACAGAAATGATCTGTTTAA 582  
 |||||  
 Db 261919 AATTTTCAATTTGATTTATTTATTTAAATTTATGATGATGATGATGATGATGATGATG 261978  
 |||||  
 Oy 583 TATAGCATATCTCATGACATATTTGACGTGACAGT 621  
 |||||  
 Db 261979 ACTATTTATATTTTATTTTATTTCCATATTTATTTATTT 262017  
 |||||

RESULT 15  
 PFA929351 347050 bp DNA linear INV 29-JAN-2003  
 LOCUS  
 DEFINITION Plasmodium falciparum strain 3D7, chromosome 5, segment 1/4.  
 ACCESSION AL929351 AL644504  
 VERSION AL929351.1 GI:23504490  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Plasmodium falciparum 3D7  
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE  
 1 Hall, N., Pain, A., Berriman, M., Churcher, C., Harris, B., Harris, D.,  
 Mungall, K., Bowman, S., Atkin, R., Baker, S., Barton, A., Brooks, K.,  
 Buckee, C. O., Burrows, C., Cherevach, I., Chillingworth, C.,  
 Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Cotton, C.,  
 Fretwell, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J.,  
 Howell, L., Goble, A., Goodhead, I., Gwilliam, R., Hamlin, N.,  
 Hance, Z., Harper, D., Hauser, H., Hornsbey, T., Holroyd, S.,  
 Horrocks, P., Humphray, S., Jagels, K., James, K. D., Johnson, D.,  
 Keshornou, A., Knights, A., Konfortov, B., Kyes, S., Larke, N.,  
 Lawson, D., Lennard, N., Line, A., Maddison, M., McLean, J., Mooney, P.,  
 Mout, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M. A.,  
 Rabbitt, E., Rajandream, M. A., Rutter, S., Rutherford, K. M.,  
 Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R.,  
 Squares, S., Stevens, K., Taylor, K., Tivey, A., Unwin, L.,  
 Whitehead, S., Woodward, J., Sultson, J. E., Craig, A., Newbold, C. and  
 Barrett, B. G.  
 Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13  
 JOURNAL Nature 419 (6906), 527-531 (2002)  
 MEDLINE 22255708  
 PUBMED 12368867  
 2 (bases 1 to 347050)  
 REFERENCE  
 AUTHORS Devlin, K., Baker, S., Davies, P., Mungall, K., Berriman, M., Pain, A.,  
 Hall, N., Bowman, S., Churcher, C., Quail, M. and Barrett, B.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-SEP-2002) P.falciparum Genome Sequencing Consortium,





gene	CDs	Query Match	Best Local Similarity	Matches	Score	DB	Length
		171	8.7%;	47.0%;	54.2;	3;	347050;
		341437	Conservative	0;	Mismatches	188;	Indels 0; Gaps 0;
		231					
		341497					
		231					
		341557					
		351					
		341617					
		411					
		341677					
		471					
		341737					
AE014837	257757 bp	AE014837	257757 bp	DNA	linear	INV	07-OCT-2002
LOCUS	Plasmodium falciparum 3D7	Plasmodium falciparum 3D7	Plasmodium falciparum 3D7	chromosome 11	section 2	of 8	of the
DEFINITION	complete sequence.	complete sequence.	complete sequence.	complete sequence.	complete sequence.	complete sequence.	complete sequence.
ACCESSION	AE014837	AE014837	AE014837	AE014837	AE014837	AE014837	AE014837
VERSION	AE014837.1	AE014837.1	AE014837.1	AE014837.1	AE014837.1	AE014837.1	AE014837.1
KEYWORDS							
SOURCE	Plasmodium falciparum 3D7	Plasmodium falciparum 3D7	Plasmodium falciparum 3D7	Plasmodium falciparum 3D7	Plasmodium falciparum 3D7	Plasmodium falciparum 3D7	Plasmodium falciparum 3D7
ORGANISM	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE	1 (bases 1 to 257757)	1 (bases 1 to 257757)	1 (bases 1 to 257757)	1 (bases 1 to 257757)	1 (bases 1 to 257757)	1 (bases 1 to 257757)	1 (bases 1 to 257757)
AUTHORS	Gardner,M.J., Hall,N.S., Fung,E., White,O., Berriman,M., Hyman,R.W., James,C., Eisen,J.A., Nelson,K.E., Bowman,S., Paulsen,I.T., Carter,C., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A., Kyte,S., Chan,M.-S., Nene,V., Shallow,S.J., Suh,B., Peterson,J., Anguilo,S., Pereira,M., Allen,J., Selengut,U., Haft,D., Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H., Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I., Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C., Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and Barrell,B.	Gardner,M.J., Hall,N.S., Fung,E., White,O., Berriman,M., Hyman,R.W., James,C., Eisen,J.A., Nelson,K.E., Bowman,S., Paulsen,I.T., Carter,C., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A., Kyte,S., Chan,M.-S., Nene,V., Shallow,S.J., Suh,B., Peterson,J., Anguilo,S., Pereira,M., Allen,J., Selengut,U., Haft,D., Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H., Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I., Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C., Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and Barrell,B.	Gardner,M.J., Hall,N.S., Fung,E., White,O., Berriman,M., Hyman,R.W., James,C., Eisen,J.A., Nelson,K.E., Bowman,S., Paulsen,I.T., Carter,C., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A., Kyte,S., Chan,M.-S., Nene,V., Shallow,S.J., Suh,B., Peterson,J., Anguilo,S., Pereira,M., Allen,J., Selengut,U., Haft,D., Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H., Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I., Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C., Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and Barrell,B.	Gardner,M.J., Hall,N.S., Fung,E., White,O., Berriman,M., Hyman,R.W., James,C., Eisen,J.A., Nelson,K.E., Bowman,S., Paulsen,I.T., Carter,C., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A., Kyte,S., Chan,M.-S., Nene,V., Shallow,S.J., Suh,B., Peterson,J., Anguilo,S., Pereira,M., Allen,J., Selengut,U., Haft,D., Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H., Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I., Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C., Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and Barrell,B.	Gardner,M.J., Hall,N.S., Fung,E., White,O., Berriman,M., Hyman,R.W., James,C., Eisen,J.A., Nelson,K.E., Bowman,S., Paulsen,I.T., Carter,C., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A., Kyte,S., Chan,M.-S., Nene,V., Shallow,S.J., Suh,B., Peterson,J., Anguilo,S., Pereira,M., Allen,J., Selengut,U., Haft,D., Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H., Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I., Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C., Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and Barrell,B.	Gardner,M.J., Hall,N.S., Fung,E., White,O., Berriman,M., Hyman,R.W., James,C., Eisen,J.A., Nelson,K.E., Bowman,S., Paulsen,I.T., Carter,C., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A., Kyte,S., Chan,M.-S., Nene,V., Shallow,S.J., Suh,B., Peterson,J., Anguilo,S., Pereira,M., Allen,J., Selengut,U., Haft,D., Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H., Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I., Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C., Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and Barrell,B.	Gardner,M.J., Hall,N.S., Fung,E., White,O., Berriman,M., Hyman,R.W., James,C., Eisen,J.A., Nelson,K.E., Bowman,S., Paulsen,I.T., Carter,C., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A., Kyte,S., Chan,M.-S., Nene,V., Shallow,S.J., Suh,B., Peterson,J., Anguilo,S., Pereira,M., Allen,J., Selengut,U., Haft,D., Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H., Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I., Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C., Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and Barrell,B.

JOURNAL. Nature 419 (6906), 498-511 (2002)  
 PUBMED. 12368864  
 REFERENCE 2 (Pages 1 to. 257757)  
 AUTHORS Gardner, M.J.:  
 TITLE Direct Submission  
 JOURNAL Submitted (13-SEP-2002) The Institute for Genomic Research, 9712  
 Medical Center Dr., Rockville, MD 20850, USA  
 FEATURES  
 source  
 1. 257757  
 /organism="Plasmodium falciparum 3D7"  
 /mol\_type="genomic DNA"  
 /isolate="3D7"  
 /db\_xref="taxon:36329"  
 /chromosome="11"  
 257. .381  
 repeat\_region  
 /rpt\_type=tandem  
 305. .373  
 repeat\_region  
 /rpt\_type=tandem  
 763. .787  
 repeat\_region  
 /rpt\_type=tandem  
 1304. .1333  
 repeat\_region  
 /rpt\_type=tandem  
 complement (<1434. .>1994)  
 /locus\_tag="PF11\_0070"  
 /locus\_tag="PF11\_0070"  
 complement (join(1434. .1828,1910. .1994))  
 /locus\_tag="PF11\_0070"  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="PAA35659.1"  
 /db\_xref="GI:23495997"  
 /translation="MCMYVCIIECTEKKKKKKKKKLPIYIHKIDVNNKAEONPL  
 VRIAGKHLFEVKROPSFTLYVTITLNNRDYTKLHAVVAVERKSLFPYLGIKK  
 ANVIYVNNKKPKWENIKSIFSSTIATVNSMDETIPSTYTLFLFISLSTKV"  
 1872. .1902  
 repeat\_region  
 /rpt\_type=tandem  
 1926. .1960  
 repeat\_region  
 /rpt\_type=tandem  
 2199. .2240  
 repeat\_region  
 /rpt\_type=tandem  
 2384. .2542  
 repeat\_region  
 /rpt\_type=tandem  
 2582. .2650  
 repeat\_region  
 /rpt\_type=tandem  
 complement (<2647. .>4074)  
 /locus\_tag="PF11\_0071"  
 /locus\_tag="PF11\_0071"  
 complement (<2647. .4074)  
 /locus\_tag="PF11\_0071"  
 /codon\_start=1  
 /product="RuvB DNA helicase, putative"  
 /protein\_id="PAA35660.1"  
 /db\_xref="GI:23495998"  
 /translation="MEINISINITNSSNNKKERVNIHSIKGLGVNTIYLHENDISL  
 DEKSYMPFNSCGLVQSPKAREASLFLVLDIIOKLAGKCIILAGPSGSKSLAIGI  
 SRINKMPPEVPLSGSEVYSNEIKTEVILEARSHIKIKKEKLVSEGVVDVVE  
 ENNCISLNKAKOINAIITLKSVKSGSKTILRPKHGOJARFKIKIGDIYIETNG  
 HVRKLGCDNYAKEDIDEPBEYVSLPGEVYHKKREVVOQSLHDIIDANNPVGBD  
 ASVNLGSLRPKTEIETEKRLVEIKTVNKFLESGLAEIIPGVYIDEAHMLDIECSY  
 LNAISLPAPILIMEGNRICTVKGIDNIEPGIIEVDLIDRLIIIKTEPYTLKEIVQ  
 IILRLAQTEKINTIEGMVYLAKIGQSSIRFAMLLPEPSRIATLEGNITIDIKHIE  
 QAELPDAKTSARHVRVDSNKKFVN"  
 4159. .4258  
 repeat\_region  
 /rpt\_type=tandem  
 4159. .4215  
 repeat\_region  
 /rpt\_type=tandem  
 4718. .4749  
 repeat\_region  
 /rpt\_type=tandem  
 4922. .4978  
 repeat\_region  
 /rpt\_type=tandem  
 5175. .5199  
 repeat\_region

```
repeat_region      /rpt_type=tandem
                    5322..5352
/rpt_type=tandem
repeat_region      5350..5397
/rpt_type=tandem
repeat_region      5405..5438
/rpt_type=tandem
repeat_region      5531..5557
/rpt_type=tandem
gene               /rpt_type=tandem
                    <5712..>6914
/rpt_type=tandem
mrna               /locus_tag="PF11_0072"
                    join(<5712..5802,5896..6191,6311..6454,6636..>6914)
/rpt_type=tandem
                    /locus_tag="PF11_0072"
                    join(5712..5802,5896..6191,6311..6454,6636..6914)
CDS                /locus_tag="PF11_0072"
                    /codon_start=1
                    /product="hypothetical protein"
                    /protein_id="AAN35661.1"
                    /translation="MKNTKSCVTCILFPYIOIWINCFPGYNINKNLGVYPSLIRK
                    KNVPEKRRQILIKQNGEKENIFPDGENTNTIKDKTKQTKNTIDIKKEKIEI
                    NEIKKSENILERSFSFKEDITIKPELKEILITQYRKIFORHDKCGSEIITII
                    LTFKIFPLTEHKKKKDPACLRGLFKCVSKRRLLVYLGRRDREMEFKISYENIRK
                    PLPRTPEYVYNKDIKYIHFNNTKRFKNNAEKKKDKLKKKQVTKILFQGN"
repeat_region      6250..6288
/rpt_type=tandem
repeat_region      6482..6530
/rpt_type=tandem
repeat_region      6484..6533
/rpt_type=tandem
repeat_region      6913..6938
/rpt_type=tandem
repeat_region      7132..7156
/rpt_type=tandem
repeat_region      7287..7327
/rpt_type=tandem
repeat_region      7691..7726
/rpt_type=tandem
gene               /rpt_type=tandem
                    <7800..>10244
/rpt_type=tandem
mrna               /locus_tag="PF11_0073"
                    join(<7800..9712,9978..10049,10155..10162,10204..>10244)
/rpt_type=tandem
                    /locus_tag="PF11_0073"
                    join(7800..9712,9978..10049,10155..10162,10204..10244)
CDS                /locus_tag="PF11_0073"
                    /codon_start=1
                    /product="hypothetical protein"
                    /protein_id="AAN35662.1"
                    /translation="MECSFLPFFPFVYVONKEIALILCSKVLKKGHIYPRKSGHEHI
                    /db_xref="GI:23496000"
                    /translacion="MECSFLPFFPFVYVONKEIALILCSKVLKKGHIYPRKSGHEHI
                    YSVUNNNVNLKKNKNSKLDRIKRIYEGVHNKKKIYELPLIKNNKNDVYIEDH
                    KDKNIDNNNNKKDKKDIDNNNDKKNKINDTPAONITNDKRYKSNPFYIENNLLK
                    LKNLLEVPSSYINKSDVITVPLNHYHDLKHLLEITPENNIINSLDYFLFKNNYNNK
                    SDININNNNNNNNNNNKNI CFPYNNIYNI IFFIYNDQDPLISSYTPSEFTNKK
                    DNI IHEKEQOLVHVI OKFLNDHYNLKDQEHKKYININFLFTKNDILDKDKDIEE
                    NMLYYPKKNILNKETNYVTNTNTNTNTKODONL IKVHNNDTFPEKRYSDPLKN
                    LTNENYVTFNPLSSYKDI LKEVIYIADITIPNNNVIKYININIKKKIHI EILLN
                    NFKKCKLNKINSHNVGRIKKQYSYLHITILSPLLEKYSNLEILKKYIKVQVKEK
                    IOKIKINSDLIFQFKQJINHDLQFQYNSKFI LPSNRYQAQGSYITTYIKIKL
                    IQLNQTANEIINYYIKGLVYKNSFNSFSKRYTLINVTIITYIENLKNKVNINF
                    NLYSPNAFESSYKDISPKKDIMITITSEMDQVKIYIQISLVVPTPKRCS"
repeat_region      8855..8890
/rpt_type=tandem
repeat_region      9936..9975
/rpt_type=tandem
repeat_region      10060..10092
/rpt_type=tandem
repeat_region      10372..10409
/rpt_type=tandem
repeat_region      10454..10485
/rpt_type=tandem
gene               /rpt_type=tandem
                    complement(<10677..>13503)
/rpt_type=tandem
                    /locus_tag="PF11_0074"
                    complement(join(<10677..10818,10922..10993,11120..11199,
```

```
CDS
11295..11397,11559..12736,12847..>13503))
/locus_tag="PF11_0074"
complement(join(10677..10818,10922..10993,11120..11199,
```

LOCUS AX598988 1501 bp DNA linear PAT 14-FEB-2003  
DEFINITION Sequence 328 from Patent WO02077272.  
ACCESSION AX598988  
VERSION AX598988.1 GI:28399128  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCES  
1  
Bellin, K., Braun, A., Distler, J., Guetig, D., Howe, A., Mueller, J.,  
Olek, A., Piepenbrock, C., Adorjan, P., Grabs, G., Lesche, R., Leu, E.,  
Lewin, A., Lipscher, E., Maier, S., Model, F., Mueller, V., Otto, T.,  
Petel, C. and Ziebart, H.  
Methods and nucleic acids for the analysis of hematopoietic cell  
proliferative disorders  
Patent: WO 02077272-A 328 03-OCT-2002;  
JOURNAL Epigenomics AG (DE)  
FEATURES  
source Location/Qualifiers  
1..1501  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"  
BASE COUNT 604 a 0 c 130 g 767 t  
ORIGIN  
Query Match 8.6%; Score 53.6; DB 6; Length 1501;  
Best Local Similarity 47.7%; Pred. No. 3;  
Matches 187; Conservative 0; Mismatches 204; Indels 1; Gaps 1;  
QY 131 TTGCTATTATTCAGCGGTAATGAGAGCTATTTGAAACAATTTCTACTATATTTTA 190  
DB 669 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 728  
QY 191 CAGATATTCGCTTGAGGTTACAGACAGCAGATATGCTTTATTTATTAAGTG 250  
DB 729 TTGTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 788  
QY 251 CTCAGAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 310  
DB 789 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 848  
QY 311 TTTCATTAAGTGAAGAGTTGTCATTAATTAATTAATTAATTAATTAATTAATTAATTA 370  
DB 849 TTTTAAATGTTGTAATGTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 908  
QY 371 GTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 430  
DB 909 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 967  
QY 431 GACTTTTACGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 490  
DB 968 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1027  
QY 491 AAGGAGAGTTTCAAACTTTTACCATTAAT 522  
DB 1028 AATGTAATTTTAAATTTTAAATTTTAAATTTT 1059  
RESULT 18  
BD092427 945 bp DNA linear PAT 27-AUG-2002  
LOCUS BD092427/c  
DEFINITION Identification of polynucleotides encoding novel helicobacter  
polypeptides in the helicobacter genome.  
ACCESSION BD092427  
VERSION BD092427.1 GI:22638038  
KEYWORDS JP 2001527393-A/208.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 945)  
Kleanthous, H., Garawi, A.A., Miller, C., Tomb, J.F. and Oomen, R.F.  
TITLE Identification of polynucleotides encoding novel helicobacter

JOURNAL polypeptides in the helicobacter genome  
Patent: JP 2001527393-A 208 25-DEC-2001;  
MERIEUX ORAVAX SOCIETE EN NOM COLLECTIF PASTEUR MERIEUX SERUMS ET  
VACCINS AGROBIOLOGICAL RESOURCES MINISTREY O SA, HUMAN GENOME  
SCIENCES INC  
COMMENT  
PN JP 2001527393-A/208  
PD 25-DEC-2001  
PF 01-APR-1998 JP 1998541947  
PR 01-APR-1997 US 08/833457,24-JUN-1997 US 08/881227 PR  
29-JUL-1997 US 08/902615  
PI HAROLD KLEANTHOUS, AMAL AL GARAWI, CHARLES MILLER, JEAN FRANCOIS  
PI TOMB,  
PI RAYMOND PETER OOMEN  
PC A01N43/04,A61K31/70  
CC Strandedness: Single;  
CC Topology: Linear;  
FH key Location/Qualifiers  
FEATURES  
source Location/Qualifiers  
1..945  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
BASE COUNT 298 a 186 c 169 g 292 t  
ORIGIN  
Query Match 8.5%; Score 53; DB 6; Length 945;  
Best Local Similarity 46.9%; Pred. No. 4.4;  
Matches 201; Conservative 0; Mismatches 225; Indels 3; Gaps 1;  
QY 19 AAAACAGAAAGCAACCCGAAACGACGTCAGAGAGCTGTTGAAGGAGTGTCTCT 78  
DB 845 AAAACCAAGCTCCCTAGCGGAAATCCAAAGCCAGAGAAAGCAATGTCCT 786  
QY 79 AAATCAGAGAGGTTACTTAAGCATGACTGACAGAGGATGCGGCTGCTATT 138  
DB 785 AAGAGCATGAGAGTGTGGGGGTTTGGGTTATTTGCGGCTAATTAATTTTGT 726  
QY 139 TATTCAGCGTAATGAGAGCTATTTGAAACAATTTCTACTAATTTTACAGAAATCA 198  
DB 725 TTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 669  
QY 199 TTTCGTTTGAAGTTACAGACAGTCAGTATGCTTATTTATTTATTTATTTATTTATTT 258  
DB 668 TTTTCCCTAGATTTCAATTAAGAAAGGCTTCAAGAGCTTTTAAACCACTGGTAAAGAC 609  
QY 259 ATAGCTATTTTATTAATGCAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTT 318  
DB 608 ACTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 549  
QY 319 CGTGTACAGAGTGTGATTAATGACATTAACAAAGGTTTAAATTAATTAATTAATTAATTA 378  
DB 548 GTCTTGAAATTTGGCTGGCTCTTTGCCCCCTAAGATTAATTAATTTTCTAAATATC 489  
QY 379 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 438  
DB 488 AACCCTATCAATGCGTCAAAACCTTTTCTTTAAAGGCTCTTGATGAGAGTTTG 429  
QY 439 CGTAGTTTA 447  
DB 428 ATCAGCTTA 420  
RESULT 19  
AE001502 10525 bp DNA linear BCT 20-JAN-1999  
LOCUS AE001502  
DEFINITION Helicobacter pylori, strain J99 section 63 of 132 of the complete  
genome.  
ACCESSION AE001502 AE001439  
VERSION AE001502.1 GI:4155264  
KEYWORDS  
SOURCE Helicobacter pylori J99  
ORGANISM Helicobacter pylori J99  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;



CDS

complement(6351..8681)  
 /gene="spot"  
 /note="similar to H. pylori 26695 gene HP0775"  
 /codon\_start=1  
 /transl\_table=11  
 /product="GUANOSINE-3',5'-BIS (DIPHOSPHATE)  
 3',-PYROPHOSPHOHYDROLASE"  
 /protein\_id="AAD06287.1"  
 /db\_xref="GI:4155268"  
 /translation="MNEIDKSDVIGFLRLIVIKVYKPKGIEVRLTLDTPKIKEN  
 ALNATSKHQYKSGEPYIVHPICASVAPFGCGEAMCAALLHDVDTCEIE  
 TIEREGODVANLVDALFKITIRKEBELGVSSODPBMWVSALTEKRLISAIQDPRL  
 VKLISDRILHNLVDLPHEQKQVIRSKETLAVAPISRLMSSIKHELEKSPYIY  
 PEYKNIKEYFLKHKQSLILKLNAFASLEKLEKPDGSGSHDFLTVRVKSPYIYK  
 MORGAVANIDEIIDLAIIRILKNPIDCYKVLGIHLNFPDIYSPFDYIAIPKENG  
 KTIHTTIFDESSVEVOIRTPDMHGAEGVSAHWKYKAGVDHDEHHEHGRMTQNK  
 YHSDDLKNDPEPEYELANDYREDIYVFSHGTYTLPVGAILDPRVYHSDLGK  
 ATDVAIINSKALNQELRSQDVVKIKGDKVIPRIMWQDKTSKAKNHLTIQRNRL  
 KEIDYKSMINILATFPGRSYFEDADLDQYKNEBKLDGQVETLTETMKSFENLAKI  
 TEETENKVFSLKEDALILEYQMSLMTGLRLGRTNVLNLTENROCKELSHFSV  
 CSSNALEIKQVLLDDCCYKYGDEILAIVTLQDKPKAIHHKPKCKKLAEDVAKVPMV  
 YIEMHKDRITVYKMPFYLGKKSVLASILITPLANEENIYGVSYLGKDKKSSHCYS  
 PEIATDKADMRALINKYQDRIVELBSLDAYELQ"

gene  
 complement(8658..8892)  
 /gene="jhp0713"  
 complement(8668..8892)  
 /gene="jhp0713"  
 /note="similar to H. pylori 26695 gene HP0776"  
 /codon\_start=1  
 /transl\_table=11  
 /product="putative"  
 /protein\_id="AAD06288.1"  
 /db\_xref="GI:4155269"  
 /translation="MKKERTSIVAAQALKNIGNRVMDNLVFAVYQVLAAGTLYN  
 MDPKRLVDVAIREIABGKIDRIDERN"

CDS  
 complement(8934..9656)  
 /gene="pyrH"  
 /note="synonym: jhp0714"  
 complement(8934..9656)  
 /gene="pyrH"  
 /note="similar to H. pylori 26695 gene HP0777"  
 /codon\_start=1  
 /transl\_table=11  
 /product="URIDYLATE KINASE"  
 /protein\_id="AAD06289.1"

gene  
 complement(8934..9656)  
 /gene="pyrH"  
 /note="synonym: jhp0714"  
 complement(8934..9656)  
 /gene="pyrH"  
 /note="similar to H. pylori 26695 gene HP0777"  
 /codon\_start=1  
 /transl\_table=11  
 /product="URIDYLATE KINASE"  
 /protein\_id="AAD06289.1"

Query Match 8.5%; Score 53; DB 1; Length 10525;  
 Best Local Similarity 46.9%; Pred. No. 2.3;  
 Matches 201; Conservative 0; Mismatches 225; Indels 3; Gaps 1;

19 AAACAGAGAAAGAACCCGAAACGAGCGTGAAGAGCTGCTGAAAGGAGTCCCT 78  
 687 AAGACCGAATCCCTAGCGGCAAAAATATCCAAAAGCAGAGAAAGGCAATGCTT 746  
 79 AAATCAGAGAGGTACTTAAGCATTTGACTACTGACAGAGGATGCTGGGCTGTATT 138  
 747 AAAAGCATGGAAGTGGGGGTTTATGCGTTATGCTGGGCTAATGATTTTGT 806  
 139 TATTCAGGCGTAATGAGACGTCATTTGAAACAATTTCTACTATTTTATGCAATCA 198  
 807 TTTTATATATGTTGGTGGATGCTTTAGCAGATGTAATCCCATGTGTTAAGATTTC 866  
 199 TTTGGTTGAGTTTACAGCAGTCAGTATATGCTTATTTATTTATGTTGCTCAGAG 258  
 867 TCCC--TAGATTTTGAAGAAAGAGCTTAAAGCTGTTTAAACAGCTGGCTAAAGAC 923  
 259 ATAGCTATTTATGATGCAATATTAATTTTATGCTGTATGCGATGATTTCAATTA 318  
 924 ACTTTTATATGCTTTGGCTGTTTATATCAATTTTAAATGAGTGGGCTTTTGTCAAT 983  
 319 CGGTACAGTGTGTGATTTAGCATCAAAAGGTTTAAATTTAAATGAGTAATTT 378  
 984 GTCTTGAATTTGGCTGCTCTTTGGCCCTTAAAGTCAATGAGCCTTAATTTCTAAATC 1043

379 AATATAAAGAGGTGAAGAGATGTTGCTTCTCAACAACACTGTGTCAGCTTTA 438  
 1044 AACCTATCAATGCGCGTCAAAACCTTTTCTTTAAAGAAATCCTGATGAGATTG 1103

439 CGTAGTTA 447  
 1104 ATCACTTTA 1112

RESULT 20  
 AE000589 10860 bp DNA linear BCT 06-APR-1999  
 LOCUS Helicobacter pylori 26695 section 67 of 134 of the complete genome.  
 DEFINITION AE000589 AE000511  
 ACCESSION AE000589.1 GI:2313895  
 VERSION  
 KEYWORDS  
 ORGANISM  
 Helicobacter pylori 26695  
 Helicobacter pylori 26695  
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;  
 Helicobacteraceae; Helicobacter.

REFERENCE  
 1 (bases 1 to 10860)  
 Tomb,J.-F., White,O., Kertlavage,A.R., Clayton,R.A., Sutton,G.G.,  
 Fleischmann,R.D., Ketchum,K.A., Klenk,H.P., Gill,S.,  
 Dougherty,B.A., Nelson,K., Quackenbush,J., Zhou,L., Kirkness,E.F.,  
 Peterson,S., Loftus,B., Richardson,D., Dodson,R., Khalak,H.G.,  
 Glodek,A., McKenney,K., Fitzgerald,L.M., Lee,N., Adams,M.D.,  
 Hickey,E.K., Berg,D.E., Gocayne,J.D., Utterback,T.R.,  
 Peterson,J.D., Kelley,J.M., Karp,P.D., Smith,H.O., Fraser,C.M. and  
 Venter,J.C.  
 The complete genome sequence of the gastric pathogen Helicobacter  
 pylori  
 JOURNAL Nature 388 (6642), 539-547 (1997)  
 MEDLINE 97394467  
 PubMed 9252185

2 (bases 1 to 10860)  
 Tomb,J.-F., White,O., Kertlavage,A.R., Clayton,R.A., Sutton,G.G.,  
 Fleischmann,R.D., Ketchum,K.A., Klenk,H.P., Gill,S.,  
 Dougherty,B.A., Nelson,K., Quackenbush,J., Zhou,L., Kirkness,E.F.,  
 Peterson,S., Loftus,B., Richardson,D., Dodson,R., Khalak,H.G.,  
 Glodek,A., McKenney,K., Fitzgerald,L.M., Lee,N., Adams,M.D.,  
 Hickey,E.K., Berg,D.E., Gocayne,J.D., Utterback,T.R.,  
 Peterson,J.D., Kelley,J.M., Cotton,M.D., Weidman,J.M., Fujii,C.,  
 Bowman,C., Wathey,L., Wallin,E., Hayes,W.S., Borodovsky,M.,  
 Karp,P.D., Smith,H.O., Fraser,C.M. and Venter,J.C.  
 Direct Submission  
 Submitted (06-AUG-1997) The Institute for Genomic Research, 9712  
 Medical Center Dr, Rockville, MD 20850, USA  
 3 (bases 1 to 10860)  
 White,O.  
 Direct Submission  
 Submitted (17-MAR-1999) The Institute for Genomic Research, 9712  
 Medical Center Dr, Rockville, MD 20850, USA

FEATURES  
 source  
 1..10860  
 /organism="Helicobacter pylori 26695"  
 /mol\_type="genomic DNA"  
 /strain="26695"  
 /db\_xref="taxon:85962"  
 complement(73..147)  
 /gene="HP0767"  
 complement(73..147)  
 /gene="HP0767"  
 /note="hypothetical protein; identified by Genemark;  
 putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="H. pylori predicted coding region HP0767"  
 /protein\_id="AAD07827.1"  
 /db\_xref="GI:2313906"  
 /translation="MEKVLAFSPSPFISFSGGGYLV"  
 complement(251..1216)  
 /gene="HP0768"

gene



Db 1917 AAAACCGACTCCCTAGCGCCGCAAAAAATCCAAAAAGCCAGAGAAAGCAAGTGCCT 1976

Qy 79 AAATCAGAGAGGTTACTTAAGCATCTGACGAGAGGAGTCTGGGCTTGCATTT 138

Db 1977 AAGAGCATGGAAGTGTGGGGGTTTGGGGTTATGGCCGGCTAATAGATTTTGT 2036

Qy 139 TATTCAGGCGTAATGGACGTCATTTTGAAACAATTTTCTATATTTTACAGATCA 198

Db 2037 TTTTATATATGTTGGTGGATGCGCTTTCGCAAAATGATCCCATGTCGTAAGA--T 2093

Qy 199 TTTGCGTTGAGTTACAGCAGTCATATATGCTTATTTATTTATTTATGTTGCACAG 258

Db 2094 TTTTCCCTAGATTTCACTAAGAAAGCGTCAAGCTGTTTACCAACTGGCTTAAGAC 2153

Qy 259 ATACCTATTTTATGATGCCAATATCTTTTATTTGCTGTACGCGATGATTTTCATTA 318

Db 2154 ACTTTTATTTATGCTTTTACCGATTTTATCATTTTATAGTGGTGGCGCTTTTATCTAT 2213

Qy 319 CGGTACAGTTGGTGCATTTATGACATCAAAAGTTTAAATTTAAATGAGTAATTT 378

Db 2214 GCTTTCGCAATTTGGCTGGCTCTTGGCCCTTAAAGCATGAGCTTAATTTTCTPAAATC 2273

Qy 379 AATATAAATAAGGTTGAAGGAATGTTGCTTCTCAACAACCTTGTTCGACTTTA 438

Db 2274 AACCTATCAATGGCGCTCAAAACCTTTTCTTTTAAAAAGCTCTTGATGGAGATTG 2333

Qy 439 CGTAGTTTA 447

Db 2334 ATCACTTA 2342

RESULT 21  
LOCUS AX345303 7167 bp DNA linear PAT 01-FEB-2002  
DEFINITION Sequence 374 from Patent WO0200928.  
ACCESSION AX345303  
VERSION AX345303.1 GI:18493189  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Olek, A., Piepenbrock, C. and Berlin, K.  
TITLE Diagnosis of diseases associated with the immune system  
JOURNAL Patent: WO 0200928-A 374 03-JAN-2002;  
EpiGenomics AG (DE)  
FEATURES  
Source location/Qualifiers  
1..7167  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"  
BASE COUNT 2279 a 47 c 1050 g 3791 t  
ORIGIN

Query Match 8.5%; Score 52.8; DB 6; Length 7167;  
Best Local Similarity 46.3%; Pred. No. 2.7;  
Matches 174; Conservative 0; Mismatches 202; Indels 0; Gaps 0;

Qy 234 TTTATTTATTTATGTTGCTCAAGAGATGCTATTTTATGATGCCAATATTACTTTTAT 293

Db 709 TATATTTGTTAGTTTATGAGATTTTAAAGTATAGATTATTAATTTTAAATTTTAA 768

Qy 294 TGCCTTACGGCAGATTCATTAACGTGATCAAGCTGTGTCATTTAGCATACAAAGT 353

Db 769 AGATTTTAGTGTATGATTTTATATATTTTAAATGTTTAAAGGTTTAAAGT 828

Qy 354 TTTTAAATTTAAATGAGTAATTTTAAATATATATAAAGGTTGAAGGAATGTTGCTC 413

Db 829 ATTAATGTTTAAAGTAATTAATGTTTAAATTTTAAAGGTTTGAATGTTTAAATTT 888

Qy 414 TCAACAAACACTTGTGACCTTTAGCTAGTTTAAAGTAAATGTTTAAAGGTAATTTGT 473

Db 889 TGTGTTAATAATGTTTAAATGTTTAAAGGTTTTCGATGTAATTTGTGTTAAGTTT 948

Qy 474 TCCATATATGATTTAATAAGAGAGATTTCAACTTTTACCATATATATGCAAGTCC 533

Db 949 ATATTTGGGTTTTTATAGCATTTTAAAGTTTATTAATTTTAAAGGTTTTTATGATA 1008

Qy 534 TTCAGGTGTGGCAGATTTATGCTTAAATACAGAAATGACTGTTTATATACGCTAT 593

Db 1009 TTTTGAAGTTTAAAGAAATTTATATTTTATGATTAATTTATGTTTTCGATTAAT 1068

Qy 594 TCCATATGCAATTTAT 609

Db 1069 TTTTAAAGAGTTT 1084

RESULT 22  
LOCUS AE010546 11258 bp DNA linear BCT 25-MAR-2002  
DEFINITION Fusobacterium nucleatum subsp. nucleatum ATCC 25586, section 88 of 197 of the complete genome.  
ACCESSION AE010546 AE009951  
VERSION AE010546.1 GI:19713801  
KEYWORDS  
SOURCE Fusobacterium nucleatum subsp. nucleatum ATCC 25586  
ORGANISM Fusobacterium nucleatum subsp. nucleatum ATCC 25586  
REFERENCE 1 (bases 1 to 11258)  
AUTHORS Kaparat, V., Anderson, I., Ivanova, N., Reznik, G., Los, T., Lykidis, A., Bhattacharyya, A., Battman, A., Gardner, W., Grechkin, G., Zhu, L., Vasileva, O., Chu, L., Kogan, Y., Chaga, O., Goldsman, E., Bernal, A., Larsen, N., D'Souza, M., Malunas, T., Pusch, G., Haselkorn, R., Fonstein, M., Kyriades, N. and Overbeek, R.  
TITLE Genome sequence and analysis of the oral bacterium Fusobacterium nucleatum strain ATCC 25586  
JOURNAL J. Bacteriol. 184 (7), 2005-2018 (2002)  
MEDLINE 21886394  
PUBMED 11889109

REFERENCE 2 (bases 1 to 11258)  
AUTHORS Kaparat, V., Anderson, I., Ivanova, N., Reznik, G., Los, T., Lykidis, A., Bhattacharyya, A., Battman, A., Gardner, W., Grechkin, G., Zhu, L., Chu, L., Kogan, Y., Chaga, O., Goldsman, E., Bernal, A., Larsen, N., D'Souza, M., Malunas, T., Pusch, G., Haselkorn, R., Fonstein, M., Kyriades, N. and Overbeek, R.  
TITLE Direct Submission  
JOURNAL Submitted (13-FEB-2002) Integrated Genomics, 2201 W. Campbell Park Drive, Chicago, IL 60612, USA  
FEATURES  
Source location/Qualifiers  
1..11258  
/organism="Fusobacterium nucleatum subsp. nucleatum ATCC 25586"  
/mol\_type="genomic DNA"  
/strain="ATCC 25586"  
/sub\_species="nucleatum"  
/db\_xref="ATCC:25586"  
/db\_xref="taxon:190304"  
complement (582..920)  
/gene="FN0337"  
complement (582..920)  
/gene="FN0337"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AAB34541.1"  
/db\_xref="GI:19713802"  
/translation="MKKNDMKRFQNAWMEYVNAKIDLVRDVNTNGIFPAIEFFENT KOIKRVFTLVPSIEIQKNVILEKFSVQNNNNOSLLTILILEINQVADKRSVINIS NEIDNNNTLN"

gene  
CDS  
1169..1513  
/gene="FN0338"  
1169..1513  
/gene="FN0338"  
/codon\_start=1

gene  
CDS  
/transl\_table=11  
/product="Hypothetical protein"  
/protein\_id="AAL94542.1"  
/db\_xref="GI:19713803"  
/translacion="MGSNFEKINQKVELGDLVLEISEKDVKAFAIYIEPFIEMI  
LGYIADOLGSEMSQIFSGFLGVGFIPLALYORPFAKKTIDBEIKVAIVEKIDPN  
ACTINASCDFP"  
1531..2004  
/gene="FN0339"  
1531..2004  
/gene="FN0339"  
/codon\_start=1  
/transl\_table=11  
/product="Hypothetical protein"  
/protein\_id="AAL94543.1"  
/db\_xref="GI:19713804"  
/translacion="MGWGSYSISLEKKELEKLNKDPDIESVETLIDIKSWQAIQ  
YLGGDICEIEGFLGVVPMVAENALDCESEGVFIYITBOIKKAYDALPFLTKEDIL  
EKVNSFEMLEDEVPIVVEDDEKEFFDIYSVLEIKKPYKNIKELALIFYS"  
complement(2084..3394)  
/gene="FN0340"  
complement(2084..3394)  
/gene="FN0340"  
/codon\_start=1  
/transl\_table=11  
/product="Hypothetical membrane-spanning Protein"  
/protein\_id="AAL94544.1"  
/db\_xref="GI:19713805"  
/translacion="MENKKYPTSVIIRKELACSLIGIFLFPVPTLNGKSTIPLDHIV  
FLKIPYKEVYGLVILIGVFLPYKKTWNKTSIVSLLKILALPFLPILFNNQ  
PEFLNKDVIPIFNKIVIPVTIVPGSIFLSLISYGLMEFVPMRPIMKPIKMT  
PGRSLDAVASPVSXSLALITNRYKGGKVTNKEAVIATGSPVSATPMVIAKTI  
LDLMNMLYFRLIYIVPLVTAIRARIPIPNKSDAIFENQKVEKDIPEKKXVA  
FNEGEVETNGSSILENVIINLKSGIMAFNIGBSLAIIGLIVLANHTPIEDIGY  
LVPEPTLISGEPEPLTAKALAGITEMPLPAVITKLSFEVKMLVAITCVSEVILFS  
ASIPCMATDIPISFKDYLIIFERFVVISIVSIPLIYLVKVM"  
complement(3499..4827)  
/gene="FN0341"  
complement(3499..4827)  
/gene="FN0341"  
/codon\_start=1  
/transl\_table=11  
/product="transport protein"  
/protein\_id="AAL94545.1"  
/db\_xref="GI:19713806"  
/translacion="MILNPVIVSVIWMVLCCLKLVNLALFISALVAGVTAGPIN  
DIMVGLITMGNGETALSYILGLTAVAINSTGVSIVSKIASVAVGKKVLLVI  
AFACFSQNLIPVHIAFIPILPPLKIMNDKLRAMACSLTFGLKAPYITLPGV  
GAIFHSIITAGETNNGMTVAQGDWKSFWIIGLPMIIGLILAFVSNKDRYKDLPL  
IGIERVADKMKAKMIALILASAFITIOALATVIRIKDKGSLPLGAVVALLMVLIF  
GVYRKKNLDERFINGVGLMGLIAFTMLVAAGNIIROTGAINELVDSIHGILGSKA  
IGVSWMLIGLITWIGISFGTIPVVAIYIPICIKGVSSSIVVLAALALGDA  
GSPADSTLGPSPSGADGQHDHMDTCVPTFTHNILLIGGFIGMFL"  
complement(4868..5371)  
/gene="FN0342"  
complement(4868..5371)  
/gene="FN0342"  
/EC\_number="5.2.1.8"  
/codon\_start=1  
/transl\_table=11  
/product="Peptidyl-prolyl cis-trans isomerase"  
/protein\_id="AAL94546.1"  
/db\_xref="GI:19713807"  
/translacion="MSLOAIITKNGEINLNFSDVAPPTVNLFTTAKGYVNGIKF  
HRVIVFMVQGDPTGTAGGPGYOGDEFKGVFNKKGLAMNAGPNTNGSQFPI  
THVPEMLNFKHTIFGEVVSQKQDVVDISIKQDPTMNEVTIIGDTDRILIDNKERYTO  
LKNPLKI"  
complement(5425..6099)  
/gene="FN0343"  
complement(5425..6099)  
/gene="FN0343"  
/codon\_start=1  
/transl\_table=11

gene  
CDS  
/product="Hypothetical protein"  
/protein\_id="AAL94547.1"  
/db\_xref="GI:19713808"  
/translacion="MEIEKMDIKIKIEDPIKRYAKEVCLQNLFLADNIKVDIKNODNLY  
EVENIEKEVLSAYENIYLSLDEEPLNLNLYENKKAFOLEPTIEKMKDNLKQYIK  
TOIKRREKLKNSGAEEVKEPFYKIKELKIKODLOKLNKLEKMKNLKDSNAIL  
OEVEOLETIDQLQPIRAEPNLSIQDKYQKEIETENKLLKKMYEIEYGTNKEILL  
KAVNSQ"  
complement(6068..7207)  
/gene="FN0344"  
complement(6068..7207)  
/gene="FN0344"  
/EC\_number="2.1.1.-"  
/codon\_start=1  
/transl\_table=11  
/product="Methyltransferase"  
/protein\_id="AAL94548.1"  
/db\_xref="GI:19713809"  
/translacion="MIFLASTTMTGESIVKDECLALGPNIKVFDGRVFEFGDFKDLI  
KANTYLRCSRVFKMAEPKALTYEELFONIKSINMODFIDEGEPFISWVSYSKYL  
YSKSDIORISKKAIVELKERYKREITLSENALYSIKIOSKODFIYWLDSGSLTK  
RGYARQKRVAPIKETTLAALVYLSKMWADVLDPGCGTGTIAEAMAIANLIPGAN  
RNFAERKMSIIEKNLMTDIRDEAFNSBLSKELTIYASDIDERSIEIAKENSEAGVE  
DDIIFEVADFNKIESPAKGYVMIIVPYGEIMGEEDIEELYRDFGNCKKGLAKMSY  
YIIRSYDEFEKAPFKKATKRNKLYNGIKCYYYQYFGDRKNKYKN"  
complement(7207..8202)  
/gene="FN0345"  
complement(7207..8202)  
/gene="FN0345"  
/codon\_start=1  
/transl\_table=11  
/product="Permease"  
/protein\_id="AAL94549.1"  
/db\_xref="GI:19713810"  
/translacion="MDMKQDQEMTINSIFAILLEPIVLYIKKKSKINDILASLSI  
VFIVIFILSLIYIPETISIKLNDIYPIISKVIITGNIANIYAKNIIYINMD  
EIDNVFTNFSVNTNTITKLLSSIGSLIDWTIGTFLAFVLAFLILDKHLIKT  
LENNIKIIPGVKNTPYIMKLSKODIFLGVSGKIYSTIVGCVYVLLITGPVVA  
ALSAILIGIGMTIPVGSITIGGIIILVAPIKTIIILVAITISQNLQVITGPYK  
ITGNKVGINTFTWITVSMTIFGNLFGIYGMFLGAFIMSLIKLFPYDLKKRAGGGE"  
complement(8321..8713)  
/gene="FN0346"  
complement(8321..8713)  
/gene="FN0346"  
/codon\_start=1  
/transl\_table=11  
/product="Hypothetical protein"  
/protein\_id="AAL94550.1"  
/db\_xref="GI:19713811"  
/translacion="MKKTKYFKIDLVYVFLIFESLILKIGSFKDVKGAKAEIIV  
DGEIKVYPILOKEKKNIFVETNIGGCVOPEDNNVRVYTSNSPLKIANQGFISPG  
VITIPDRLVKVVYDSESDSGIDPFVAR"  
complement(8700..9602)  
/gene="FN0347"  
complement(8700..9602)  
/gene="FN0347"  
/EC\_number="4.1.1.65"  
/codon\_start=1  
/transl\_table=11  
/product="Phosphatidylserine decarboxylase"  
/protein\_id="AAL94551.1"  
/db\_xref="GI:19713812"  
/translacion="MNFEEKIOYIRKNGEIKTEKVMGSGALFLIYNPFGKALHAV  
KRKLSDMYGRKMSKPSKERTKFSVEEMEDNMEYKRPIDVYASFNDFPBEIKDGA  
RKIILYNENIVYSPADGKLAYQNKREVDKFPVSKSTLLEFPNDKELAKKEDDTFV  
IVRLAPADVYHFRPVDGEISEIKILGYYVSSTHAKITFRFICENKREVALILTE  
KFQDIAMEDIQATVWGIVQYKTNSSVYKGEERGVFLGSGSTCILVEKKNKVIDKD  
IIENTONKIEFRIYGEKFGNEKN"  
complement(9664..11169)  
/gene="FN0348"  
complement(9664..11169)  
8.5%; Score 52.8; DB 1; Length 11258;

Query Match



Best Local Similarity 46.7%; Pred. No. 2.4;  
Matches 168; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

```
OY 161 ATTTGGAACATTTTCTACTATATTTTACAGATCATTTGGTTGAGTTACAGAC 220
    |||||
DB 7839 ATTTGTTATATTTGTACTATTTTACTTACAAATTTGTAATAATATTTCTATTTCT 7898
    |||||
OY 221 AGTCAGTATATGCTTATTTATTTATGTTGCTCAAGATAGCTATTTTATGATGCCAA 280
    |||||
DB 7899 CCATTTTATGTTATTAATAATTTTGTGCAATATGTTGCTATATTTTCCCTGTTA 7958
    |||||
OY 281 TATTACTTTTATGCTGTTACGCGATGATTCATTCAGTTCACAGTTGGTGATAT 340
    |||||
DB 7959 TTATCACTTTTTCGAAATTTGAGATTAATATCATTTTAATGTTTATGATGAATTA 8018
    |||||
OY 341 GGACTCAAGATTTTATTTTAAATTTAATGAGTAATTTATATATAAAGGTTGAAG 400
    |||||
DB 8019 TTTCTGGAATTAATCTATCATGATTAATAATTAATTAATTAATTAATTAATTAATTA 8078
    |||||
OY 401 GAATGTTTCTCTCAACAAACATTTGCACTTTTACGTTAGTTAGTTCAAGTAATG 460
    |||||
DB 8079 AACTATATGCAATATATCATTTATTTACTTTTCTTTTAAATATTAATTAATTAATGCT 8138
    |||||
OY 461 TTATAGGATTTGTTCCATATATGATTAATAAAGAGAGTTTCAACTTTTACCATAT 520
    |||||
DB 8139 CTAAAGATATGCTATAAATATTTGACATTAATGATTAATAAATATTTTCCATTTAT 8138
    |||||
```

## RESULT 23

AP002091

LOCUS AP002091 106763 bp DNA linear PRI 30-OCT-2001  
DEFINITION Homo sapiens genomic DNA, chromosome 6q25.2, clone: CITD-2058112,  
complete sequence.

ACCESSION

AP002091

AP002091.2 GI:15208256

VERSION

KEYWORDS

SOURCE

Homo sapiens (human)  
Homo sapiensEukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

Shimizu, N. and Asakawa, S.

TITLE

Homo sapiens DNA chromosome 6 SEQUENCE

JOURNAL

Published Only in Database (2000)

REFERENCE

2 (bases 1 to 106763)

AUTHORS

Shimizu, N. and Asakawa, S.

TITLE

Direct Submission

JOURNAL

Submitted (22-May-2000) Nobuyoshi Shimizu, Keio University, School  
of Medicine, Molecular Biology, 35 Shinanomachi, Shinjuku-ku, Tokyo  
160-8582, Japan (E-mail: nshimizu@med.keio.ac.jp,  
Tel:81-3-3351-2370, Fax:81-3-3351-2370)

On Aug 16, 2001 this sequence version replaced gi:8096474.  
Location/Qualifiers  
1..106763

COMMENT

FEATURES

Source

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="6"

/map="6q25.2"

/clone="CITD-2058112"

/clone\_lib="CIT Approved Human Genomic Sperm Library D"

148..499

/evidence=not experimental

/rpt\_family="THE1A"

repeat\_region

843..1125

/evidence=not experimental

/rpt\_family="AluSg"

complement(2227..2493)

/evidence=not experimental

/rpt\_family="AluSx"

repeat\_region

2713..3024

/evidence=not experimental

/rpt\_family="AluSg"

repeat\_region

3027..3133

/evidence=not experimental

/rpt\_family="MLT1G1"

3465..3499

/evidence=not experimental

/rpt\_family="AT\_rich"

complement(3706..4131)

/evidence=not experimental

/rpt\_family="MSTB"

complement(4260..4558)

/evidence=not experimental

/rpt\_family="AluY"

4847..5046

/evidence=not experimental

/rpt\_family="MIR"

5587..5886

/evidence=not experimental

/rpt\_family="AluSc"

complement(6350..6557)

/evidence=not experimental

/rpt\_family="L1ME"

complement(6640..6750)

/evidence=not experimental

/rpt\_family="L1ME"

6751..7043

/evidence=not experimental

/rpt\_family="AluSg"

complement(7044..7075)

/evidence=not experimental

/rpt\_family="L1ME"

7085..7108

/evidence=not experimental

/rpt\_family="TTTTG.n"

complement(7109..7388)

/evidence=not experimental

/rpt\_family="AluY"

complement(7389..7827)

/evidence=not experimental

/rpt\_family="L1ME"

8049..8251

/evidence=not experimental

/rpt\_family="MSTA"

8281..8517

/evidence=not experimental

/rpt\_family="MSTB"

8752..8784

/evidence=not experimental

/rpt\_family="AT\_rich"

9041..9209

/evidence=not experimental

/rpt\_family="MIR"

9211..9244

/evidence=not experimental

/rpt\_family="TC.n"

9244..9271

/evidence=not experimental

/rpt\_family="TTA.n"

10750..10830

/evidence=not experimental

/rpt\_family="MIR"

11137..11175

/evidence=not experimental

/rpt\_family="AT\_rich"

11602..11908

/evidence=not experimental

/rpt\_family="AluSg1"

13292..13577

/evidence=not experimental

/rpt\_family="AluJo"

complement(14159..14311)

/evidence=not experimental

/rpt\_family="MER5B"

15530..15605

/evidence=not experimental

```

/evidence=not experimental
/rpt_family="Ct-rich"
complement(15609..15872)
repeat_region
repeat_region
/evidence=not experimental
/rpt_family="L2"
1630f..16537
/evidence=not experimental
/rpt_family="MIR"
17545..17847
/evidence=not experimental
/rpt_family="AluSx"
18093..18126
/evidence=not experimental
/rpt_family="Al-rich"
18129..18416
/evidence=not experimental
/rpt_family="AluYb8"
18457..18506
/evidence=not experimental
/rpt_family="Al-rich"
18773..18875
/evidence=not experimental
/rpt_family="L1PA16"
18935..19236
/evidence=not experimental
/rpt_family="AluSx"
19244..19267
/evidence=not experimental
/rpt_family="(GGGA)n"
19277..19603
/evidence=not experimental
/rpt_family="L1PA16"
19604..20394
/evidence=not experimental
/rpt_family="LTR1"
20395..21725
/evidence=not experimental
/rpt_family="L1PA16"
21969..21992
/evidence=not experimental
/rpt_family="(TA)n"
22191..22311
/evidence=not experimental
/rpt_family="(TA)n"
22311..22311
/evidence=not experimental
complement(22313..22608)
repeat_region
repeat_region
/evidence=not experimental
/rpt_family="AluSx"
23055..23098
/evidence=not experimental
/rpt_family="(T)n"
23693..23693
/evidence=not experimental
/rpt_family="AluSg"
25106..25411
/evidence=not experimental
/rpt_family="AluY"
26220..26220
/evidence=not experimental
/rpt_family="MIR"
27269..27291
/evidence=not experimental
/rpt_family="Al-rich"
29229..29229
complement(28687..29229)
/evidence=not experimental
/rpt_family="MER76"
29695..30982
/evidence=not experimental
/rpt_family="L1MC4"
30984..31037
/evidence=not experimental
/rpt_family="(TA)n"
3115..31235
/evidence=not experimental

```

---

```

/rpt_family="L1MC4"
31450..31514
/evidence=not experimental
/rpt_family="(TCTA)n"
31659..31693
/evidence=not experimental
/rpt_family="Al-rich"
31900..32060
/evidence=not experimental
/rpt_family="(TA)n"
32061..32188
repeat_region

```

Query Match 8.5% Score 52.8; DB 9; Length 106763;  
Best Local Similarity 45.8%; Pred. No. 1.3; Mismatches 217; Indels 0; Gaps 0;  
Matches 183; Conservative 0;

```

186 TTTTACAGAAATCATTCGTTGAGTTGAGTTACAGACAGCAGTATATGCTTTATTTATTTA 245
32630 TATTATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 32689
246 TGTGCTCAAGAGTATGCTATTTTATGATGCCATATATACCTTTATTTGCTGTAACGC 305
32690 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 32749
306 ATGATTTTCATTCAGTGTACAGAGTTGTCATTTATGACATCAAAAGTTTAAATTTAA 365
32750 ATATATAGTATATATATATATATATATATATATATATATATATATATATATATATAT 32809
366 ATGAGTAAATTTATATATATATATATATATATATATATATATATATATATATATATAT 425
32810 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 32869
426 TGTTCGACTTTTACGTATTTAGTTCAAGTATTTCTTATAGTATGTTCCATATATAT 485
32870 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 32929
486 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 545
32930 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 32989
546 AGATTATATGCTTATATATACAGAAATGATCTGTTTATAT 585
32990 AAATTTTACATTTATATATATATATATATATATATATATATATATATATATATATAT 33029

```

RESULT 24  
PFA929358 254050 bp DNA linear INV 29-JAN-2003  
LOCUS Plasmodium falciparum strain 3D7, chromosome 9; segment 4/5.  
DEFINITION Al929358 AL844508  
ACCESSION AL929358.1 GI:23505147  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Plasmodium falciparum 3D7  
Plasmodium falciparum 3D7  
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
1 (bases 1 to 254050)  
Hall, N., Pain, A., Bertman, M., Churcher, C., Harris, B., Harris, D.,  
Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooks, K.,  
Buckee, C. O., Burrows, C., Cherevach, I., Chillingworth, C.,  
Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corton, C.,  
Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J.,  
Feltwell, T., Goble, A., Goodhead, I., Gilliam, R., Hamlin, N.,  
Hancock, Z., Harper, D., Hauser, H., Hornsby, T., Holroyd, S.,  
Horrocks, P., Humphrey, S., Jagels, K., James, K. D., Johnson, D.,  
Kerhornou, A., Knights, A., Kontorov, B., Kyes, S., Laite, N.,  
Lawson, D., Lennard, N., Line, A., Maddison, M., McLean, J., Mooney, P.,  
Moule, S., Murphy, L., Oliver, K., Omond, D., Price, C., Quail, M. A.,  
Rabindrnwiesch, E., Rajandream, M. A., Rutter, S., Rutherford, K. M.,  
Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R.,  
Squares, S., Stevens, K., Taylor, K., Tivey, A., Unwin, L.,  
Whitehead, S., Woodward, J., Wiston, J. E., Craig, A., Newbold, C. and  
Barrell, B. G.

TITLE  
Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13

JOURNAL Nature 419 (6906), 527-531 (2002)  
MEDLINE 22255708  
PUBMED 12368867  
REFERENCE 2 (bases 1 to 254050)  
AUTHORS Sanders, M., Hauser, H., Baker, S., Unwin, L., Mungall, K., Berriman, M., Pain, A., Hall, N., Bowman, S., Churcher, C., Quail, M. and Barrett, B.  
TITLE Direct Submission  
Submitted (20-SEP-2002) P.falciparum Genome Sequencing Consortium, The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK  
COMMENT For more information about this sequence or the Malaria Project, see [http://www.sanger.ac.uk/Projects/P\\_falciparum](http://www.sanger.ac.uk/Projects/P_falciparum).  
FEATURES  
Source Location/Qualifiers  
1. 254050  
/organism="Plasmodium falciparum 3D7"  
/mol\_type="genomic DNA"  
/isolate="3D7"  
/db\_xref="taxon:36329"  
/chromosome="9"  
763. 2532  
/gene="PF11215w"  
763. 2532  
/gene="PF11215w"  
/note="Moderately similar to Mus musculus splicing factor 3a, subunit 3, 60kd TR:AAH09141 (EMBL:BC009141) (501 aa) fasta scores: E(): 1.7e-24, 30.756% id in 595 aa. No significant pfam hit; SMART hit to SMO0355, zinc finger/ProfileScan hit to P850171, Zn-finger (C2H2) in matrin and other RNA-binding proteins"  
/codon\_start=1  
/product="spliceosome-associated protein, putative"  
/protein\_id="CAD51929.1"  
/db\_xref="GI:23505148"  
/translation="MSOFLRQIRYRHEIRIVIRKAIKALIDKYNKKKILVDYALNYLVEKIOEKQLLEYINDEDKLKKEMMLSGTGGEDKADIKNRYEKYLRDHRKINIKIETIKSSLYKLEALKCSKYQSFSPLEKKGKYVDLNFYDFIMKKIKERIISFKKKKEMASNNRRKKTGDNNNNNNNNNNTTANINEFEMDYVTFYHNTFRFYIPRYCKRYNEKYSYENLNIYMIISFEKINIVADVCTYKEAYDFYENKKEIKGMEKYTDLDYECNINNNKLVASGEGYOSLYCKYADNDIKYLNKKYTVGLEIEIKNIEKEDNNLARCEVLIQLYANILNKIQITTONIKQKAYTTIDETQKKYKKEBEDERKRRKEVHLISLMDNNNELNNQNLLOLENNNISISMSDSDSDMSISNNEQOKDLDEENEDNDNNKPIYNPLNPLGFNDNKPYPWLYKLGSKSEYKCEICGNYSYGSA TFEKHGFYEWRSFGMKCKIKIPNTHLPKETIKIDALNIYELKKKETQWNIFFKPOEYVCEDSKGVNNIKAYDILKROGL"  
join(4370..4451,4555..4604,4733..4964,5141..5186,5417..5444,5735..5773,5790..5825)  
/gene="PF1120w"  
/gene="PF1120w"  
join(4370..4451,4555..4604,4733..4964,5141..5186,5417..5444,5735..5773,5790..5825)  
/note="1 probable transmembrane helix predicted for PF11220w by TMHMM2.0 at aa 74-96"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAD51930.1"  
/db\_xref="GI:23505149"  
/translation="METFNHIDIKKYREDLLVLRNTNNVLFKMKQVDSIACLSK EOKICGLTISGNNINENYTYTIVTKKHITTYIYVVFCLGICISFTIYVLIIS CQVLTNRHMLRIKYSNNKLVQVRLKIKKKKSTQRPFCFLFPHNHFVIFSNFKYV MREPLINMC1"  
join(6108..6182,6392..6547,6648..6699,6851..6903,7009..7062,7151..7204,7343..7424,7603..7710,7795..7853)  
/gene="PF11225w"  
join(6108..6182,6392..6547,6648..6699,6851..6903,7009..7062,7151..7204,7343..7424,7603..7710,7795..7853)  
/gene="PF11225w"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAD51931.1"  
/db\_xref="GI:23505150"  
/translation="MNEKKGKIGSVHEKQNLSTHQKEDLLIDYVCDLKEEKRRS TEKTKGDFVFKSVDRNKADYAVRNNRKMHLRLDEDEPITDILINENKIGTVR ITNPLPIERDELDKNYMKWIKTERLMRPNERLEVEQMEKYLESIKILTHIDKQV

gene  
CDS  
KLAKAIKSHYPRGIVGDSIYENTVLVKQHEIRKKKYEIKERLKEGEVLEKNG KNGNLFTEENN"  
complement(8468..9508)  
/gene="PF11230c"  
complement(8468..9508)  
/gene="PF11230c"  
/note="asparagine-rich  
Signal anchor predicted for PF11230c by SignalP 2.0 HMM (Signal peptide probability 0.001, signal anchor probability 0.995) with cleavage site probability 0.001 between residues 27 and 28  
4 probable transmembrane helices predicted for PF11230c by TMHMM2.0 at aa 12-34, 38-56, 63-82 and 97-116"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAD51932.1"  
/db\_xref="GI:23505151"  
/translation="MEERKWSYIMVSYIFLYCPEFPLYSYEIIFLEPQTKIDYVFC LFTFSLSFIYFCSGKGLYICTPLITISLIEIYEYNNIPKDYDEFLKFLVRL FSAVNLVDFLYLHINKLNKIKDEKLNDDKOSTIOSVAPPTTENNHNHNNNNNN IHNHNHNNHNNHNNHNNNDPCNNNDPCNNNYKQNDPLKENTYVSHNNFKEN IPHVHGTGITSKTLPOENIYUAKHSIPIFEKIIINKDINISHMTEMQNDSPFNH QVNNLSSNNINDKITEPITTHNLNKNNAVYTHIIONNHVPNTYEHPRGLQYYPG NFHTHNNGNINIO"  
join(11425..12434,12559..12655,12774..12908)  
/gene="PF11235w"  
join(11425..12434,12559..12655,12774..12908)  
/gene="PF11235w"  
/note="No pfam hit; ScanProsite hit to P500430; InterPro hit to Ton-B dependant receptor protein"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAD51933.1"  
/db\_xref="GI:23505152"  
/translation="WTDINNNKSKLVKKNKKNKNNKNNKNNKNTLLVSENDYKKE KKKKKHKNLQVYIHNHNNNNHDKHIDENQSNHLYVNNRTRKKKNELIDKWL NKKKKKKTKKKTKNNNTIYDNNININDEEKKKKVFPYISVIRENVANNKI IHIDNEVAKTYADALENNNKOKYKKKKKSENPEDIVNSFLRYNEVWYTNNAVIO KLNQTRVNNIYHOGYKOKKKKNNHVSITIKHLYKYPNNKNTIADGCEAATART LDGYTISFPLDIQNNHYTCNTQLPLNNSYDCLFSLSLANTWDPKITFESVRL KKGATVLIAEVSRFTYVKAIFKMNNVGFLSKINIMDDFFVFFPKQNEVDISS INERKIRKVSLSLPPCIYKRR"  
complement(113154..14893)  
/gene="PF11240c"  
complement(113154..14893)  
/gene="PF11240c"  
/EC\_number="6.1.1.15"  
/note="Similarities to prolyl-t-RNA synthases of bacterial origin; N-terminal signal peptide. Similar to Lactococcus lactis prolyl-tRNA synthetase pros or 112127 SWALL:Q9CDDT4 (EMBL:AE006441) (616 aa) fasta scores: E(): 1.6e-09, 35.21% id in 213 aa, and to Escherichia coli prolyl-tRNA synthetase pros or dtpa or b0194 SWALL:GYP\_ECOLI (SWALL:P16659) (572 aa) fasta scores: E(): 7.3e-08, 31.91% id in 235 aa. Pfam match to PF00587, tRNA synthetase class II (G, H, P, S and T)  
Signal peptide predicted for PF11240c by SignalP 2.0 HMM (Signal peptide probability 0.720, signal anchor probability 0.018) with cleavage site probability 0.591 between residues 18 and 19"  
/codon\_start=1  
/product="prolyl-t-RNA synthase, putative"  
/protein\_id="CAD51934.1"  
/db\_xref="GI:23505153"  
/translation="MLIKIILILGLFHHVILSLKLVNKKRSNVFVNKKNLKNYP KKKRYLSNNQSYIGSNYIFNRIFNINISNGKISSELRANIVIRDVANGIYVILP GFRVINKIIDLKNGLEKLSHMSLSIILAKKLNISDSKISYDSEFLVYKOKOK NDKSSESLKKIIEDGYILSPICESSLSLINIVENTIKCLPLIHQYVKKFKN EKPERKISFKSKFIMDGYSPHSNECLNMYEKYECCKNIPEELKLSFNITIKKK KDKNNALSHSRFVLSRDKYKREKAHIFKIGDYISNLDIKYDKKNEKKNITLMSYG IGIRLLYFLLENFYEEGIRKLPQVAVFSGVYLIQNOKRSYASTKSLINMYKDSM EKGKNSIQQNRNEDQDDCDECGSTFASNTHNDKVNKDDTLNNDTLNSNNTLNHD NGVINNTLYTSPLSITSYIILTLMLFNTFKNNKVDIYVDDTDLHLSRKLKHCCLIG

gene  
 complement(join(15848..16581,16944..17610))  
 /gene="PF11245c"  
 complement(join(15848..16581,16944..17610))  
 /gene="PF11245c"  
 /note="Plasmodium falciparum protein phosphatase-beta TR-015920 (EMBL:U89025) (466 aa). fasta scores: E(): 1.9e-180, 99.142% id in 466 aa. pfam match to PF00149, Ser/Thr protein phosphatase; SMART hit to SM00156, Protein phosphatase 2A homologues, catalytic domain."  
 /codon\_start=1  
 /product="Protein phosphatase-beta"  
 /protein\_id="CAD51935.1"  
 /db\_xref="GI:23505154"  
 /translation="MKCELSKREDLIFNNVNTKANKLKRANKVKGDLTYAEKPKDTS  
 SLIEEDDESELDCCGSEKKYKISNENINTEPNPKENYINYNLSVKELSDSE  
 EVSESEYNNKIKKNNVNNPNNNSIEIYTSINDEQNNKNNVNNVNNVNNVNNV  
 VDEMIKLLKELKIEYKLMCDLLIDLKRENCVINVAVAGIHGQFPDLLE  
 LEHIGLPPDVVYFLGIDYVDRGYSCGECGACPKIKYPSRVTLGNHESRQITK  
 VYGFDECRKRYNNNNIYWKYITVDYDYLPLAIIINDEFCGHGSLSCLOTIDINK  
 LDREKEIPQDGAICDILNSDPAGEDELTGKWRSPRGAVLPSEKTESFLRNKLS  
 CIGRAHQLVQDFQWNNHNDKVTIFSAIPYCYRCNCASLMLVDFEMKEDPTFTAP  
 LRANAQALRRNVQYML"  
 join(19483..19560,19785..19859,19939..20005,20102..20182,  
 20291..20343)  
 /gene="PF11250w"  
 join(19483..19560,19785..19859,19939..20005,20102..20182,  
 20291..20343)  
 CDS  
 Query Match 8.5%; Score 52.6; DB 3; Length 254050;  
 Best Local Similarity 44.4%; Pred. No. 1.1;  
 Matches 211; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

QY 135 TATTATTAGCGGCGATGAGCGATTTTGAACAATTTCTACTATTATTACAGA 194  
 |||||  
 Db 237780 TGTATTATTAACTTATTCATTATTATTATTATTATTATTATTATTATA 237839  
 |||||  
 QY 195 ATCATTTCGGTTGAGGTACAGCAGCAGATATGCTTATTATTATTATGTCCTCA 254  
 |||||  
 Db 237840 TATATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 237899  
 |||||  
 QY 255 AGAGATAGCTATTATTATGATGCAATATTACTTTTATGCTTACGCGATGATTC 314  
 |||||  
 Db 237900 TTTATCTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 237959  
 |||||  
 QY 315 ATTACGTGACAAAGTTGGCATTTATGACTCAAAAGCTTTTAAATTTAAATGAGTAA 374  
 |||||  
 Db 237960 ATTACTTATTTATCTTATTATTATTATTATTATTATTATTATTATTATTATT 238019  
 |||||  
 QY 375 ATTTAATATATAAAGGTTGAAAGGATGTTGGTCTCAACAAGCTTGGCAT 434  
 |||||  
 Db 238020 ATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 238079  
 |||||  
 QY 435 TTACGTAGTTAGTTCAAGTATTTGTTATAGTATGTTCCATATGATATAAAG 494  
 |||||  
 Db 238080 ATTAAATATTTATTATTATTATTAGTATGATTTTATTATTATTATTATTATA 238139  
 |||||  
 QY 495 AAGAGTTTCAACTTTTACCATTAATTATGCAAGTCTTCAGGTGGCAGATTAAT 554  
 |||||  
 Db 238140 TTATTATTATATAAATTTATTATTATTATTATTATTATTATTATTATTATTAT 238199  
 |||||  
 QY 555 GGTATATACAGATAGTACTGTTTATATACGTAATTCATGACATTAAT 609  
 |||||  
 Db 238200 TTATTATTATTAATTATTATTATTATTATTACATATTATTATTATTATTATT 238254  
 |||||

RESULT 25  
 AL390756 99003 bp DNA linear HMG 13-JUN-2001  
 LOCUS AL390756  
 DEFINITION Homo sapiens chromosome 1 clone RP11-378A4, \*\*\* SEQUENCING IN  
 PROGRESS \*\*\* 29 unordered pieces.  
 ACCESSION AL390756  
 VERSION AL390756.2 GI:9931005

# KEYWORDS SOURCE ORGANISM

HTG; HTGS, PHASE1; HTGS, CANCELLED.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

# REFERENCE AUTHORS TITLE JOURNAL

McLay, K.  
 Direct Submission  
 Submitted (12-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 requests: clone.requests@sanger.ac.uk  
 On Aug 27, 2000 this sequence version replaced gi:9581725.

# COMMENT

----- Genome Center  
 Center: Sanger Centre  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: humquery@sanger.ac.uk  
 Project Information  
 Center project name: ba378A4

----- Summary Statistics  
 Assembly program: XGAP4; version 4.5  
 Sequencing vector: plasmid; L08752; 100% of reads  
 Chemistry: Dye-terminator; BT-amer sham; 26% of reads Chemistry:  
 Dye-terminator Big Dye; 73% of reads  
 Consensus quality: 84348 bases at least Q40  
 Consensus quality: 89481 bases at least Q30  
 Consensus quality: 92139 bases at least Q20  
 Insert size: 96203; sum-of-contigs  
 Insert size: 188174; 2.5% error; agarose-fp  
 Quality coverage: 2.40x in Q20 bases; sum-of-contigs Quality  
 coverage: 1.82x in Q20 bases; agarose-fp  
 -----

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 29 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 2813: contig of 2813 bp in length  
 \* 2814 2913: gap of 100 bp  
 \* 2914 7380: contig of 4467 bp in length  
 \* 7381 7480: gap of 100 bp  
 \* 7481 11099: contig of 3619 bp in length  
 \* 11100 11199: gap of 100 bp  
 \* 11200 13897: contig of 2638 bp in length  
 \* 13898 13997: gap of 100 bp  
 \* 13998 16453: contig of 2456 bp in length  
 \* 16454 16553: gap of 100 bp  
 \* 16554 19055: contig of 2502 bp in length  
 \* 19056 19155: gap of 100 bp  
 \* 19156 21888: contig of 2733 bp in length  
 \* 21889 21988: gap of 100 bp  
 \* 21989 25542: contig of 3554 bp in length  
 \* 25543 25642: gap of 100 bp  
 \* 25643 30331: contig of 4699 bp in length  
 \* 30332 30431: gap of 100 bp  
 \* 30432 32487: contig of 2056 bp in length  
 \* 32488 32587: gap of 100 bp  
 \* 32588 36034: contig of 3447 bp in length  
 \* 36035 36134: gap of 100 bp  
 \* 36135 38214: contig of 2080 bp in length  
 \* 38215 38314: gap of 100 bp  
 \* 38315 40332: contig of 2018 bp in length  
 \* 40333 40432: gap of 100 bp  
 \* 40433 43102: contig of 2570 bp in length  
 \* 43103 43202: gap of 100 bp  
 \* 43203 48312: contig of 5110 bp in length  
 \* 48313 48312: gap of 100 bp  
 \* 48313 52476: contig of 4164 bp in length  
 \* 52477 52576: gap of 100 bp  
 \* 52577 54553: contig of 2077 bp in length







[illegible]

Query Match	8.4%;	Score 52.2;	DB 3;	Length 341050;
Best Local Similarity	46.0%;	Pred. No. 1.2;		
Matches 177;	Conservative 0;	Mismatches 208;	Indels 0;	Gaps 0;

[illegible]

QY	345	TACAAAGCTTTTAAATTTAAATGAGTAAATTTAAATATAATAAAGGTTGAAGAAAT	404
Db	23686	TGACCAAGAATCATATATATATTTGCTATTTTAAAGAAATATAATATATATATATA	236923
QY	405	GTTTGCTTCTCAACAAACACTGTTCGACTTTTAGTACTTTAGTTCAAGTAATGTTAT	464
Db	236926	TATAATATATATATACATATATTTATATGTCATTAAGAAATGTTAACTCATTTGTAATTTT	236985
QY	465	AGGATTTGTCATATATGATTTATATAAGAGAGATTTCAAACTTTTCCATTATATATA	524
Db	236986	AAGAAATATATATATATATATTTATATATAATTTAAATATATATATATCCTTTATTAATATATA	237045
QY	525	TGCAAGTCCTTCAGGTGGCGAGATTATATGCTTATATACAGAAATAGTACTGTTTTATA	584
Db	237046	TGAATTAATATAGCTTTTAAATATCATATATATAAATAATATATATATATATTTT	237105
QY	585	TACGTAATTCCTATGACATATTT	609
Db	237106	TATATCCACATATATAATATAAT	237130

RESULT 28			
AX344554	AX344554	349980 bp	DNA
LOCUS			linear
DEFINITION	Sequence 5 from Patent WO0200932.		
ACCESSION	AX344554		
VERSION	AX344554.1	GI:18492440	
KEYWORDS			
SOURCE	synthetic construct		
ORGANISM	synthetic construct		
	artificial sequences.		

REFERENCE  
1  
Olek, A., Piepenbrock, C. and Berlin, K.  
Diagnosis of known genetic parameters within the micro-  
Patent: WO 0200932-A 5 03-JAN-2002;

Epigenomics AG (DE)	Location/Qualifiers
FEATURES	1. .349980
source	

```

/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:12630"
/note="chemically treated genomic DNA (Homo sapiens)-Original length of seq 1: 3.67378 <22>-split as follows.-seq 01 0.000.001 TO 0.349.980-seq 02 0.300.001 649.980-seq 03 600.001 949.980-seq 04 900.001 1.249.980-seq 05 1.200.001 1.549.980-seq 06 1.500.001 1.849.980-seq 07 1.800.001 2.149.980-seq 08 2.100.001 2.449.980-seq 09 2.400.001 2.749.980-seq 10 2.700.001 3.049.980-seq 11 3.000.001 3.349.980-seq 12 3.300.001 3.649.980-seq 13 3.600.001 3.673.778 <22>-Original length of seq 2: 3.673778 <22>-split as follows.-seq 14 0.000.001 TO 0.349.980-seq 15 0.300.001 649.980-seq 16 600.001 949.980-seq 17 900.001 1.249.980-seq 18 1.200.001 1.549.980-seq 19 1.500.001 1.849.980-seq 20 1.800.001 2.149.980-seq 21 2.100.001 2.449.980-seq 22 2.400.001 2.749.980-seq 23 2.700.001 3.049.980-seq 24 3.000.001 3.349.980-seq 25 3.300.001 3.649.980-seq 26 3.600.001 3.673.778"
BASE COUNT      92349 a      5068 c      82248 g      170315 t
ORIGIN

```

	Query Match	8.4%	Score 52.2	DB 6	Length 349980	
	Best Local Similarity	44.4%	Pred. No. 1.2			
	Matches 210	Conservative 0	G mismatches 263	Indels 0	Gaps 0	
QY	135 TATTATTACAGCGGTATGGAGCTCATTTTGAACAATTTCTACTATATTTTACAGA	194				
Db	143314 TTTTTTTTAAAGTTGTTTTATATTATTTTGAATTTTTTGTTTTAACTTTATTACG	14333				
QY	195 ATCATTTTCGTTTGAAGTTACAGACAGTACGTATATGCTTTATTATTATATGTTGCGCA	254				
Db	143374 TTACGTTTAAAGTTTGAATTTTGAACAAGCTGTTTTATCTTTTTTTTTTTTTTTTGAGATGGA	14344				



OY	255	AGAGATAGCTATTATTATATGACCAATATACCTTTTATATGCGTTAACGCCATGAGATTTC	314
Db	143434	GTAATTTGGTAGTTGTTTATTTTAAATATATATATTAAGATGATGATTTGAGGTAAATTTT	143493
OY	315	ATTACGTGTACAAAGTTGGTGCAATTATGACCTACAAAGTTTTTAAATTTAAATGAGATA	374
Db	143494	AATATGAAATTATAGTTTTTGGTTTATTTGTTTATGCTTTTGAAATCGTAGATTAATATGATGAG	143555
OY	375	ATTATATATATATAAAGGTTGAAAGAAATGTTTGCTTCTCAACAACAATTGTGCACT	434
Db	143554	AGGAGATATTATATGAAATATGAGAGTAGAATATTTTAAATGTTGAAATATTAAGTTTAGAGA	143613
OY	435	TTTAGCTAGTTTAGTTCAAGTAAATGTTTATAGTATGTTCCATATATGATTTATPAAAGG	494
Db	143614	GGAGTGAAGTTTGAGAAAGAGTTAGGTTTATGTTTATCGGAAAGAAATTAAGATTTGGAAA	143673
OY	495	AGAGTTTCAAACTTTTACCATTAATATATATGACGAAGCTCTTCAGGTGTGGCAATTATAT	554
Db	143674	AGATATGAGTTTAAAGTTGGGTGAGTAGTTTTCGAAGGGCGGTGTTAGTAAAGATTAATTT	143733
OY	555	GCTTATATACAGAAATAGTACTGTTTATATATACGTAATTTCCATGACAACTTA	607
Db	143734	ATATTTGATGTATGAAAGAGATTTTGTGGAATATGTTTATTAAGATTAATTAATTA	143786

LOCUS	AX251268	5297 bp	DNA	linear	PAT 05-OCT-2001
DEFINITION	Sequence 236 from Patent WO0168912.				
ACCESSION	AX251268				
VERSION	AX251268.1 GI:15984691				
KEYWORDS	synthetic construct synthetic construct artificial sequences.				
ORGANISM	Olek, A., Piepenbrock, C. and Berlin, K. Diagnosis of diseases associated with tumor suppressor genes and oncogenes Patent: WO 0168912-A 236 20-SEP-2001; Epigenomics AG (DE)				
REFERENCE	1. 5297 Location/Qualifiers /organism="synthetic construct" /mol type="genomic DNA" /db xref="taxon:32630" /note="chemically treated genomic DNA (Homo sapiens)"				
JOURNAL	Source				
FEATURES	BASE COUNT 1353 a 107 c 1126 g 2711 t ORIGIN				
Query Match	8.4%; Score 52; DB 6; Length 5297; Best Local Similarity 46.6%; Pred. No. 4.2; Matches 203; Conservative 0; Mismatches 230; Indels 3; Gaps 1;				
QY	158	GTCAATTTGAACAATTTCTACATATATTTTACAGAACTATTGGTTGAGTTACAG	217		
Db	2669	GTAATTTTTCGATTTTGGTTTACGTTTATATGTAATGTAATGTTTATTTAGT	2728		
QY	218	CACAGTCAGTATA--TCCTTATTTATTTATNGTGTCAAGAGATCACTATTTATTTGA	274		
Db	2729	TATATTTGGAATAGAGAGTTTACGTTTTTTGTATGTTTTTTGAAGTATTTTGAAGTT	2786		
QY	275	TGCCAATTAATCTTTTATATGCTGTACGCGATGAGATTTCATTAAGTACAGTTGGTG	334		
Db	2789	GTATGTTTTTATTTATTTTGTATTAAGAAATTTGTTTATTTTGTGAGAGATAGA	2848		
QY	335	CATTATGACATCAAAAGTTTAAATTTAAATGAGATAATTTATATTAATAAAGGCT	394		
Db	2849	AAATTAATTTGTTTGTATTTTATTTTATTTGTTTAAATTTTGAATTAAGAGATTTT	2908		
QY	395	TGAAAGAAATGTTTGCCTTCTCAACAAACACTGTTGCACTTTTACGTAGTTAGTTCAAG	454		

Db 2909 TTAATAATGAAATAGTTTTTAGTAGAATATAGTCGGTGTTTATAGTTTAATTAATT 2968  
 Qy 455 TAAATGTTAATAGGATTTGTTCCATATAATGATTTAAAGAGAGCTTTCAAACCTTTTAC 514  
 Db 2969 TTGATATATGTTTATTTTAAATTTAATAGTTAAATTTTAAAGAGCTTTATTTTATTA 3028  
 Qy 515 CATATATTTATGCAAGCTTCAGCTGCGACAGTTATAGCTTAATACGAGATTAATAC 574  
 Db 3029 GAATTTTAGGTTGATGTTATTTATATATTTTGAATGATGATTTAATAATTTAGGGTTTTAT 3088  
 Qy 575 TTGTTTATATACGCT 590  
 Db 3089 ATTTTTTTTTTTAGGTT 3104

LOCUS	AX346461	7128 bp	DNA	linear	PAT 01-FEB-2002
DEFINITION	Sequence 1532 from Patent WO0200928.				
ACCESSION	AX346461				
VERSION	AX346461.1 GI:18494347				
KEYWORDS					
SOURCE	synthetic construct				
ORGANISM	synthetic construct				
REFERENCE	artificial sequences.				
AUTHORS	1. Olek, A., Piepenbrock, C. and Berlin, K.				
TITLE	Diagnosis of diseases associated with the immune system				
JOURNAL	Patent: WO 0200928-A 1532 03-JAN-2002;				
FEATURES	Epigenomics AG (DE)				
source	Location/Qualifiers				
	1..7128				
	/organism="synthetic construct"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32630"				
	/note="chemically treated genomic DNA (Homo sapiens)"				
BASE COUNT	1875 a 207 c 1610 g 3436 t				
ORIGIN					
Query Match	8.4%; Score 52; DB 6; Length 7128;				
Best Local Similarity	46.5%; Pred. No. 3.8;				
Matches 240; Conservative	0; Mismatches 270; Indels 6; Gaps 2;				
OY	78	TAAATCAGAAAGAGTTACTAAGACATTCAGTCTACGACGACGAGGATGCTGGGGCTTGCTAT	137		
Db	4809	TAAATTTATGATTTATTTAATTTGATTTTATAGTTTAAATGATATGATGTTAATAT	4868		
OY	138	TTATTCAGCGGTAATGGAGCGTCAATTTGAAACAATTTCTACTATATTTTTCAGATATC	197		
Db	4869	TTAGTTGATGATGATGATGATTTATTTGTTATTTTATTTTAAATATTTTATTTAAAT	4928		
OY	198	ATTTCGTTGAGGTTACAGACAGCGATATAGCTTTATTTATTTATCTGTCACAGA	257		
Db	4929	TTTAAATATGATATTAATAATGATTAATATGATTAATGATTTGATTTGTTT	4988		
OY	258	GATAGCTATTTTATGATGCCAATATTAATCTTTATTTGCTGTACGGCATGATTTTCAT	317		
Db	4989	TAT-----TATGATGATGAATAATATTTGAAAGTTTTTTTTATGAGATTAATAGTAATTT	5043		
OY	318	ACGCTGTCAGAGTGGTCATTTATGACTACAAAGGTTTTAAATTTAAATGAGTAAT	377		
Db	5044	TAGTTATATTTTATGATTTATTTATTTATTTAATAA-TTATTTATGAGTATTTTATTTATG	5102		
OY	378	TAAATATAATAAAGGCTTGAAGAAAGTATGCTTCCAAACAAACCTTGTCAGCTTT	437		
Db	5103	TTTATATTTGTTTATGATTTATTTAGGAGATATAAAGTATATATATATGTTTATTTAT	5162		
OY	438	ACGTAGTTTATGTTCAAGTAATTTGTTATAGTATTTGCCATATATGATTTTAAAGAGA	497		
Db	5163	GGGGTTTATATTTTATGATGAAATTTATGATTTTATGATTTAAATATATGAGAAATATTT	5222		
OY	498	GTTTTCAAACTTTTTCATTTATATATGCAAGTCTTCAGTGTGGCAGATATATATGCT	557		



/note="assembly\_fragment:02193  
fragment\_chain:1"  
misc\_feature  
38739..52285  
/note="assembly\_fragment:02147  
fragment\_chain:1"  
misc\_feature  
52386..66020  
/note="assembly\_fragment:00176  
fragment\_chain:2"  
misc\_feature  
66121..138641  
/note="assembly\_fragment:01542  
fragment\_chain:2"  
misc\_feature  
138742..145950  
/note="assembly\_fragment:01549  
fragment\_chain:3"  
misc\_feature  
146061..190195  
/note="assembly\_fragment:02266  
fragment\_chain:3"  
misc\_feature  
190296..193436  
/note="assembly\_fragment:00691.0"  
misc\_feature  
193537..217879  
/note="assembly\_fragment:00862.0"  
BASE COUNT 69938 a 38381 c 38675 g 70078 t 807 others  
ORIGIN

Query Match 8.4%; Score 52; DB 2; Length 217879;  
Best Local Similarity 44.6%; Pred. No. 1.5;  
Matches 205; Conservative 0; Mismatches 255; Indels 0; Gaps 0;

Oy 162 TTTTGAACATTTTCTACTATTTTACAGAAATCTTGGGTTGAGTTACAGCA 221  
Db 39578 TTTTATATATTTTACATTTTATTTTAAATATATATATATATATATAC 39519  
Oy 222 GTGAGTATGCTTATTTATTTTATGTCAGAGATGCTATTTATGATGCAAT 281  
Db 39518 TAAATATTTTATTTATTTATTTATGTTATTTATTTATTTATTTATTA 39459  
Oy 282 ATTACTTTTATGCTGTTAGCGAGATTTTACGTATACGTTAGTTCAGATT 341  
Db 39458 ATTA 39399  
Oy 342 GACATCAAGGTTTAAATTTAAATGAGTAAATTTAATATATATATATATAT 401  
Db 39398 TATTTGTAATTTTATTTAGTTACATCTTTTATTTTAAAGATTAAAT 39339  
Oy 402 AATGTTGCTCTCAACAAACCTGTCGCTTTTACGTATGTTAGTCAAGTA 461  
Db 39338 ATATTTACTTTCTTACATTAATTTGTTATTTATTTATTTATTTATTA 39279  
Oy 462 TATAGTATGCTTCATATATGATTTAAAGAGATTTTCAACCTTTTACATTA 521  
Db 39278 TA 39219  
Oy 522 TTATGCAAGTCTTCAGGTGCGCAGATTAATGCTTAAACAGGAATGACTG 581  
Db 39218 TATATATATATACACATTAATATATATATATATATATATATATAT 39159  
Oy 582 ATATAGGCTATTCCTATGACATTTATGAGTGCAGAT 621  
Db 39158 ATTTACCTAACACTCACTATTTATTTATTTATCTAAT 39119

RESULT 33  
AX345066 5523 bp DNA linear PAT 01-FEB-2002  
LOCUS  
DEFINITION Sequence 137 from Patent WO0200928.  
ACCESSION AX345066  
VERSION AX345066.1 GI:18492952  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Olek, A., Piepenbrock, C. and Berlin, K.

TITLE Diagnosis of diseases associated with the immune system  
JOURNAL Patent: WO 0200928-A 137 03-JAN-2002;  
Epidemiology AG (DE)  
FEATURES  
SOURCE  
1..5523  
Location/Qualifiers  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"  
BASE COUNT 1378 a 119 c 1366 g 2660 t  
ORIGIN

Query Match 8.3%; Score 51.8; DB 6; Length 5523;  
Best Local Similarity 45.6%; Pred. No. 4.5;  
Matches 182; Conservative 0; Mismatches 217; Indels 0; Gaps 0;

Oy 185 TTTTACAGATCTTTGCTTTAGTTACAGCACTGCTATGCTTATTTATTTAT 244  
Db 726 TTTTATATATTTTGTGTTATGTTGATTTTATGTTATATATATATATAT 785  
Oy 245 ATGTTGCTCAAGATATGCTATTTATTTATGATGCAATATTTATGCTGTTAC 304  
Db 786 ATGAGTTTATGTTTAAATATTTATGCTGTTGTTTATTTATTTTATTTAT 845  
Oy 305 CATGATTTATATGCTGTTACAGTTGCTGATTTATGACTACAAAGTTTAA 364  
Db 846 AATTAAGATGTTAGTTGATGTAAGGGTTAAGAGATCGTTATTTGTTTAT 905  
Oy 365 AATGAGTAAATTTATATATATATATATATATATATATATATATATAT 424  
Db 906 ATATATGAGGATTTATTTATTTAGTGTATGATGTTATATGTTTATTTAT 965  
Oy 425 TTGTTGCTCAAGATATGCTATTTATTTATGATGTTATTTATTTATTTAT 484  
Db 966 TTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1025  
Oy 485 TTATTAAGAGATTTTCAACCTTTTACATTTATTTATGAGTCTTCAGGTG 544  
Db 1026 GTAGTATATTTATGAGTAAATGAGTGTATGATGAAATTTATTTATGAT 1085  
Oy 545 CAGATTTATGCTTAAATACAGAAATGACTGTTTAT 583  
Db 1086 TATATTTATATATATATATATATATATATATATATATATATATAT 1124

RESULT 34  
AY242996/c AY242996 15575 bp DNA circular INV 15-MAR-2003  
LOCUS  
DEFINITION Anthereaea pernyi mitochondrial, complete genome.  
ACCESSION AY242996  
VERSION AY242996.1 GI:28974757  
KEYWORDS  
SOURCE  
ORGANISM  
mitochondrion Anthereaea pernyi (Chinese oak silkworm)  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
Bombycoidea; Saturniidae; Saturniinae; Saturniini; Anthereaea.  
REFERENCE 1 (bases 1 to 15575)  
Liu, Y., Xia, Y., Lu, C. and Xiang, Z.  
AUTHORS Complete mitochondrial genome of Anthereaea pernyi  
TITLE Unpublished  
JOURNAL 2 (bases 1 to 15575)  
Liu, Y., Xia, Y., Lu, C. and Xiang, Z.  
AUTHORS Direct Submission  
REFERENCE Submitted (23-FEB-2003) Key Sericultural Laboratory of the  
Agricultural Ministry, Southwest Agricultural University, No.216  
Tianshengqiao Road, Chongqing 400716, China  
FEATURES  
SOURCE  
1..15575  
Location/Qualifiers  
/organism="Anthereaea pernyi"  
/organism="mitochondrion"  
/mol\_type="genomic DNA"  
/strain="Zao 1"

```
tRNA
/db_xref="taxon:7119"
/dev_stage="pupa"
/county="China: Henan"
1..67
/product="tRNA-Met"
77..140
/product="tRNA-Ile"
complement(138..206)
/product="tRNA-Gln"
263..1276
/gene="ND2"
263..1276
/gene="ND2"
/codon_start=1
/transl_table=5
/product="NADH dehydrogenase subunit 2"
/protein_id="AA053273.1"
/db_xref="GI:28974758"
/translation="MYPLNSKMFPTILPSTLIAISSNMGCGWIGLEINLSFI
PLISSPFLNLAESLKYFLQSIASINFLVILKMTFFKFLNDPFSIMISSML
MKMGAPHPFMPNIVEGLSWPNFPLMTWQITPMILSYFNKFLYMTIFMII
GAIGLNOISLRKLAFSINLGMMLPAIMSENLMTFVFTMSFLSIMCFEFL
NMFPINOLFNNISPLIKINLINFSLGCLPPLGFPFKMIMINFLNNKRYLLTFL
PYMMSLIIFFYIRITTYCTCLMNTFKMKLKTSLKNNFLMNFPSLISIMGMISTE
LPL"
1285..1352
/product="tRNA-Trp"
complement(1345..1405)
/product="tRNA-Cys"
complement(1425..1490)
/product="tRNA-Tyr"
<1497..3027
/gene="COI"
<1497..3027
/gene="COI"
/notes="start codon undetermined: TAA stop codon is
completed by the addition of 3' residues to the mRNA"
/codon_start=1
/transl_except=(pos:3027,aa:TERM)
/transl_table=5
/product="cytochrome c oxidase subunit I"
/protein_id="AA053274.1"
/db_xref="GI:28974759"
/translation="RKMLYSINHKDICTYIFFGIACGAVGTSLSLLRAELGTGSL
IGDDQIYNTIYTAHAFIMLFVWMPIMIGFGFNWLLPLMLGAPDAPFPKNNWSTWLL
PSLTLISSIYENGAGTGTVPPLSSINIAHGSSVDLAIFSLHAGISILTAIN
FATTLIINMNNISFQDMPLEWAVGITAFLLLSLPVLAGATMLLDRNLTSEFD
PAGGGPILYOHLEFMPFGHREYVILLIPGGMISHIISOEGKKEGCLSMLYMMA
IGLGGIYVAHNTFTYGMIDIRAVTSAITMIIVPTGKIKFSMLATLHGQINSPS
ILMSLGFVFLFTVGGITGVILANSSIDITLHDTLYVAHFHYVLSKGAFAIMGCFIH
WYPLFTGLINPFLKIOFFIMFGVNLTFPQHFLGLGMPRRYSDDPDSLWNLLI
SLGYSILSLAVMLMLIIMESMIYORIALFSLNMBSSIEWYONLPPAHASYNEPLIL
SN"
3028..3093
/product="tRNA-Leu"
3094..3775
/gene="COII"
3094..3775
/gene="COII"
/notes="TAA stop codon is completed by the addition of 3'
residues to the mRNA"
/codon_start=1
/transl_except=(pos:3775,aa:TERM)
/transl_table=5
/product="cytochrome c oxidase subunit II"
/protein_id="AA053275.1"
/db_xref="GI:28974760"
/translation="MAYWSNLNIQNGASPLMEQIIFHDHTLILIMTILYGYMIS
LFPNSIYNFLLEGOMIELIWLIPALITLIFALPRLLYLDELNNPLITKSGH
QWWSYEDSPHNIEDPSYINPLINDINNFRLLVDNRIILPMNNQIRIMWTASDVI
HSMTIPSLGIKVDANGRLNQTNVFINRGIIFYGQSEICGANHSMPYIVESISIKN
FINMINVS"
3776..3855
tRNA
/product="tRNA-Lys"
3879..3915
/product="tRNA-Asp"
3952..4119
/gene="ATP8"
3952..4119
/gene="ATP8"
/codon_start=1
/transl_table=5
/product="ATP synthase F0 subunit 8"
/protein_id="AA053276.1"
/db_xref="GI:28974761"
/translation="MPQMPMPNMWFSFIIFIIIFLINMNYIIFNYINKNNKLIN
NKMKSFSMKW"
4113..4790
/gene="ATP6"
4113..4790
/gene="ATP6"
/codon_start=1
/transl_table=5
/product="ATP synthase F0 subunit 6"
/protein_id="AA053277.1"
/db_xref="GI:28974762"
/translation="MMGNLFSIPDPSYINILSLNMLSTPLGLLFLPYSEWFINRHF
FWKFIILPKLNEFKTLKNNYFQSGTFIFSLSPSFLPFLGLFPYIFTSGLSL
SLGISPLMSFMYXGWINNYQWHPHMIPOGTFILMPVLLIETISNIIIRPGLAV
RUTANMTAGHLMLTILSNMGSILPSYILMPLIFQIILLIESAVAVIQSYVIALST
LYSSEVN"
4790..5578
/gene="COIII"
4790..5578
/gene="COIII"
/codon_start=1
/transl_table=5
/product="cytochrome c oxidase subunit III"
/protein_id="AA053278.1"
/db_xref="GI:28974763"
/translation="MKULYSHPYHLVDYSPWPLTGAIGVMTLYTGLVKFHPNFSINL
IIGYITILTYQWRDISRRCYQGGHTLIVTGLRMGMLFTVSELPFLSPFWAF
FHSLSPNIDIGAMPPLSTPFPNPHIPLANTILLISGTYVAHHAALMENNLTOT
NQSLIFLIGIYFTILOAVEYLAPSIADSIYGAIFPMATGFHGLHVIITGLPLIV
CFIRHLSHFSSTHFGFEAAAWYHFDVVMFLPLYSIYMWGN"
5581..5646
/product="tRNA-Gly"
5647..6000
/gene="ND3"
5647..6000
/gene="ND3"
/codon_start=1
/transl_table=5
/product="NADH dehydrogenase subunit 3"
/protein_id="AA053279.1"
/db_xref="GI:28974764"
/translation="MIIMMLIMITLILANIMFISILSKSPFDRKSCPFEGCFD
PKSSARIPSLHPFLITITLIFDVEIALPIIYLFXTVNFPMWTKTSFFIILLLL
GVYHNNOMLWTIN"
5999..6064
/product="tRNA-Ala"
6064..6128
/product="tRNA-Arg"
6129..6193
/product="tRNA-Asn"
6195..6260
/product="tRNA-Ser"
6262..6327
/product="tRNA-Glu"
complement(6338..6408)
/product="tRNA-Phe"
complement(6394..8149)
/gene="ND5"
complement(6394..8149)
/notes="TAA stop codon is completed by the addition of 3'
```

```
tRNA
/product="tRNA-Lys"
3879..3915
/product="tRNA-Asp"
3952..4119
/gene="ATP8"
3952..4119
/gene="ATP8"
/codon_start=1
/transl_table=5
/product="ATP synthase F0 subunit 8"
/protein_id="AA053276.1"
/db_xref="GI:28974761"
/translation="MPQMPMPNMWFSFIIFIIIFLINMNYIIFNYINKNNKLIN
NKMKSFSMKW"
4113..4790
/gene="ATP6"
4113..4790
/gene="ATP6"
/codon_start=1
/transl_table=5
/product="ATP synthase F0 subunit 6"
/protein_id="AA053277.1"
/db_xref="GI:28974762"
/translation="MMGNLFSIPDPSYINILSLNMLSTPLGLLFLPYSEWFINRHF
FWKFIILPKLNEFKTLKNNYFQSGTFIFSLSPSFLPFLGLFPYIFTSGLSL
SLGISPLMSFMYXGWINNYQWHPHMIPOGTFILMPVLLIETISNIIIRPGLAV
RUTANMTAGHLMLTILSNMGSILPSYILMPLIFQIILLIESAVAVIQSYVIALST
LYSSEVN"
4790..5578
/gene="COIII"
4790..5578
/gene="COIII"
/codon_start=1
/transl_table=5
/product="cytochrome c oxidase subunit III"
/protein_id="AA053278.1"
/db_xref="GI:28974763"
/translation="MKULYSHPYHLVDYSPWPLTGAIGVMTLYTGLVKFHPNFSINL
IIGYITILTYQWRDISRRCYQGGHTLIVTGLRMGMLFTVSELPFLSPFWAF
FHSLSPNIDIGAMPPLSTPFPNPHIPLANTILLISGTYVAHHAALMENNLTOT
NQSLIFLIGIYFTILOAVEYLAPSIADSIYGAIFPMATGFHGLHVIITGLPLIV
CFIRHLSHFSSTHFGFEAAAWYHFDVVMFLPLYSIYMWGN"
5581..5646
/product="tRNA-Gly"
5647..6000
/gene="ND3"
5647..6000
/gene="ND3"
/codon_start=1
/transl_table=5
/product="NADH dehydrogenase subunit 3"
/protein_id="AA053279.1"
/db_xref="GI:28974764"
/translation="MIIMMLIMITLILANIMFISILSKSPFDRKSCPFEGCFD
PKSSARIPSLHPFLITITLIFDVEIALPIIYLFXTVNFPMWTKTSFFIILLLL
GVYHNNOMLWTIN"
5999..6064
/product="tRNA-Ala"
6064..6128
/product="tRNA-Arg"
6129..6193
/product="tRNA-Asn"
6195..6260
/product="tRNA-Ser"
6262..6327
/product="tRNA-Glu"
complement(6338..6408)
/product="tRNA-Phe"
complement(6394..8149)
/gene="ND5"
complement(6394..8149)
/notes="TAA stop codon is completed by the addition of 3'
```

tRNA

residues to the mRNA"  
 /codon\_start=1  
 /transl\_except=(pos:complement(6394),aa:TERM)  
 /transl\_table=5  
 /product="NADH dehydrogenase subunit 5"  
 /protein\_id="AA053280.1"  
 /db\_xref="GI:28974765"  
 /translation="MFEKNYSVCFISFIFLPMFSSNFIIMLIIIFIMNDLIYFEMEII  
 SLNSVSVMSIILDMVMSILFMFEVCLISVVIYYSKSYMOSELNDRIFILVLFVS  
 MLIISIPNNVSIILGMDGLVSYCLVITYONKSYNAGMELTRNIGDVLIVMI  
 SMNLNYSNNYIFELFPMKNDPMFMEVIGWIIILAAKRSQIPSSSLPAPMAAPTV  
 SALVHSSTLVTVGYELIRNMLIDMFRRILILSLITFMAGISANYPELKKIT  
 ALSTLSQGLMSILSMGPPDLAFPHLLTAMFAMLMFCAGAVIHMMNDIODIRWG  
 GISMFIPMTSLCINISNMALCGIPFLAGFYSKDILLEMVSFNILVFFPYISTGL  
 TMFVTEFLILYTIMDENLAVIYNLYDEDFIMKSMVLLFMSVIGSFLSPMMIESYP  
 YMIYLPNNKMGWIVYVSFGALRGYISNNNIYSMKFELLTYKLSNPFCLMFMPEPIS  
 TVGVTEFLNLGONLMKNIIDLGSEMSISGGAMDIMNYSIFPNIIFPMNKLILBSF  
 ILMFPIMLILFLYLSNL"  
 /complement(8150..8218)  
 /product="tRNA-His"  
 /complement(8228..9568)  
 /gene="tRNA"  
 /complement(8228..9568)  
 /gene="tRNA"  
 /codon\_start=1  
 /transl\_table=5  
 /product="NADH dehydrogenase subunit 4"  
 /protein\_id="AA053281.1"

Query Match 8.3%; Score 51.8; DB 3; Length 15575;

Best Local Similarity 46.0%; Pred. No. 3.4;  
 Matches 214; Conservative 0; Mismatches 247; Indels 4; Gaps 1;

145 GGGCTAATGGAGTCATTTTGAACAATTTTCTACTATATTTTAAAGATCATTTCGG 204  
 |||||  
 14179 GGGGGAAGATTAATAAAGAAATTAATAAATAAATTTTAAATATAATATTTTA 14120  
 |||||  
 205 TTGAGGTTACAGACAGTCAGTATATCTTATTTTATTTATGTTGCTCAAGATAGCT 264  
 |||||  
 14119 ATGGGGTTAAGATTTTAAATAGAAAATTTTATTTATTTATGTAATAGTAATGT 14060  
 |||||  
 265 ATTTTATGATGCCAATATTACTTTTATGCTGCTTAAAGGATGATTTTCAATCGTGA 324  
 |||||  
 14059 AAAAGATTTTGAATAATTTGAAAAAATTTTAAAAAGTAATAATTTATTTGTA 14000  
 |||||  
 325 CAAGTGGTGCAATTATGCAACAAGTTTAAATTTAAATGAGTAATTTAAATA 384  
 |||||  
 13999 TCTTGTATCAAGATTTTAAAAATTTTAAATTTTAAATTTCTCGAATTTTAAAG 13940  
 |||||  
 385 ATAAAAGGTTGAAGAAATGTTGCTTCAACAACAACCTGTTGCACTTTTACGTAGT 444  
 |||||  
 13939 AGTTAATTAATTAATAAGTATTGTCATTAATTTTAAATTAATTAATTAATGA 13880  
 |||||  
 445 TTAGTTCAAGTAATTTGTAAGGATTTGTCATATATGATTAATAAAGAGAGTTTCA 504  
 |||||  
 13879 AAGTTAATGCTTTTAAATATATCTAGTTT---TTTGAAGAAAAATTTAAATTTTAA 13824  
 |||||  
 505 AACTTTTACCATTAATTTATGCAAGCTTCAGAGTGTGCAATTTATGCTTAATACA 564  
 |||||  
 13823 TTTAATTTAAATAAATAATTAATTTTAAATTTTAAATTTAAATAATTAATTAAG 13764  
 |||||  
 565 GGAATAGTACTGTTTATATAGCTTAATCTATGACATTAAT 609  
 |||||  
 13763 GGAATAGCTTTGAATTAATTTTATATGATTAATTAATTAATTT 13719

RESULT 35  
 AX347026 73334 bp DNA linear PAT 01-FEB-2002  
 LOCUS AX347026  
 DEFINITION Sequence 2097 from Patent WO0200928.  
 ACCESSION AX347026  
 VERSION AX347026.1 GI:18494912  
 KEYWORDS

SOURCE  
 ORGANISM synthetic construct  
 synthetic construct  
 artificial sequences.  
 REFERENCE  
 1 Olek.A., Piepenbrock,C. and Berlin,K.  
 Authors Diagnosis of diseases associated with the immune system  
 TITLE Patent: WO 0200928-A 2097 03-JAN-2002;  
 JOURNAL Epigenomics AG (DE)  
 FEATURES  
 source  
 1..73334  
 /organism="synthetic construct"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32630"  
 /note="chemically treated genomic DNA (Homo sapiens)"  
 BASE COUNT 18968 a 858 c 15329 g 38179 t  
 ORIGIN

Query Match 8.3%; Score 51.8; DB 6; Length 73334;

Best Local Similarity 49.8%; Pred. No. 2.2;  
 Matches 131; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

264 TATTTATGATGCCAATATTACTTTTATGCTGTTACGCGATGATTCATTACGTGT 323  
 |||||  
 57666 TAAATATGAGGCTTAATGTTAGTTTATTTTGTGAGATTTTAAATTTTAAATTTTAA 57725  
 |||||  
 324 ACAAGTGGTGCATTAATGACCTCAAAAGTTTTTAAATTTAAATGAGTAATTTAAT 383  
 |||||  
 57726 TAAATATGAGGCTTAATGTTAGTTTATTTTGTGAGATTTTAAATTTTAAATTTTAA 57785  
 |||||  
 384 AATPAAAAGGTTGAAGAAATGTTGCTTCAACAACAACCTGTTGACCTTTTACGTAG 443  
 |||||  
 57786 TTTAGTAATTAATAAATAGTTTATGTTTATTTAAGAAAGAAATTTTGTAAAT 57845  
 |||||  
 444 TTTAGTTCAAGTAATTTGTTATAGTATGTTCCATATATGATTTAATAAAGAGAGTTTC 503  
 |||||  
 57846 TTAAGTGAAGTAATTAATAATTTTAAATTTTAAATTTTATGATTAATTTTAAAGGGGCTTT 57905  
 |||||  
 504 AACTTTTACCATTAATTTATG 526  
 |||||  
 57906 AATATTTTGTGATTAAGGTTATG 57928  
 |||||

RESULT 36  
 AX356493 73334 bp DNA linear PAT 06-FEB-2002  
 LOCUS AX356493  
 DEFINITION Sequence 127 from Patent WO0181622.  
 ACCESSION AX356493  
 VERSION AX356493.1 GI:18620983  
 KEYWORDS  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 artificial sequences.  
 REFERENCE  
 1 Olek.A., Piepenbrock,C. and Berlin,K.  
 Authors Diagnosis of diseases associated with dna repair  
 TITLE Patent: WO 0181622-A 127 01-NOV-2001;  
 JOURNAL Epigenomics AG (DE)  
 FEATURES  
 source  
 1..73334  
 /organism="synthetic construct"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32630"  
 /note="chemically treated genomic DNA (Homo sapiens)"  
 BASE COUNT 18968 a 858 c 15329 g 38179 t  
 ORIGIN

Query Match 8.3%; Score 51.8; DB 6; Length 73334;

Best Local Similarity 49.8%; Pred. No. 2.2;  
 Matches 131; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

264 TATTTATGATGCCAATATTACTTTTATGCTGTTACGCGAGATTCATTACGTGT 323  
 |||||  
 57666 TAAATATGAGGCTTAATGATTTTATTTATGATTTTAAATTTTAAATTTTAAATTTTAA 57725  
 |||||



```
* 4409 5165: contig of 757 bp in length
* 5166 5265: gap of 100 bp
* 5266 6031: contig of 766 bp in length
* 6032 6131: gap of 100 bp
* 6132 6896: contig of 765 bp in length
* 6897 6996: gap of 100 bp
* 6997 7771: contig of 775 bp in length
* 7772 7872 8651: contig of 780 bp in length
* 8652 8752 9524: contig of 773 bp in length
* 9525 9625 10447: contig of 823 bp in length
* 10448 10547: gap of 100 bp
* 10548 11406: contig of 859 bp in length
* 11407 11507 12265: contig of 759 bp in length
* 12266 12365: gap of 100 bp
* 12366 13122: contig of 757 bp in length
* 13123 13223 13980: contig of 758 bp in length
* 13981 14080: gap of 100 bp
* 14081 14922: contig of 842 bp in length
* 14923 15022: gap of 100 bp
* 15023 15805: contig of 783 bp in length
* 15806 15905: gap of 100 bp
* 15906 16664: contig of 759 bp in length
* 16665 16765 17537: contig of 773 bp in length
* 17538 17637: gap of 100 bp
* 17638 18409: contig of 772 bp in length
* 18410 18509: gap of 100 bp
* 18510 19278: contig of 769 bp in length
* 19279 19378: gap of 100 bp
* 19379 20137: contig of 759 bp in length
* 20138 20237: gap of 100 bp
* 20238 21012: contig of 775 bp in length
* 21013 21112: gap of 100 bp
* 21113 21892: contig of 780 bp in length
* 21893 21992: gap of 100 bp
* 21993 22810: contig of 818 bp in length
* 22811 22910: gap of 100 bp
* 22911 23712: contig of 802 bp in length
* 23713 23812: gap of 100 bp
* 23813 24594: contig of 782 bp in length
* 24595 24694: gap of 100 bp
* 24695 25452: contig of 758 bp in length
* 25453 25552: gap of 100 bp
* 25553 26318: contig of 766 bp in length
* 26319 26418: gap of 100 bp
* 26419 27200: contig of 782 bp in length
* 27201 27300: gap of 100 bp
* 27301 28095: contig of 795 bp in length
* 28096 28195: gap of 100 bp
* 28196 28977: contig of 782 bp in length
* 28978 29077: gap of 100 bp
* 29078 29883: contig of 806 bp in length
* 29884 29983: gap of 100 bp
* 29984 30776: contig of 793 bp in length
* 30777 30876: gap of 100 bp
* 30877 31661: contig of 785 bp in length
* 31662 31761: gap of 100 bp
* 31762 32517: contig of 756 bp in length
* 32518 32617: gap of 100 bp
* 32618 33392: contig of 775 bp in length
* 33393 33492: gap of 100 bp
* 33493 34254: contig of 762 bp in length
* 34255 34354: gap of 100 bp
* 34355 35148: contig of 794 bp in length
* 35149 35248: gap of 100 bp
* 35249 36040: contig of 792 bp in length
* 36041 36140: gap of 100 bp
* 36141 36912: contig of 772 bp in length
```

```
* 36913 37012: gap of 100 bp
* 37013 37807: contig of 795 bp in length
* 37808 37907: gap of 100 bp
* 37908 38664: contig of 757 bp in length
* 38665 38765: gap of 100 bp
* 38765 39602: contig of 838 bp in length
* 39603 39702: gap of 100 bp
* 39703 40477: contig of 775 bp in length
* 40478 40577: gap of 100 bp
* 40578 41343: contig of 766 bp in length
* 41344 41443: gap of 100 bp
* 41444 42203: contig of 760 bp in length
* 42204 42303: gap of 100 bp
* 42304 43068: contig of 764 bp in length
* 43068 43167: gap of 100 bp
* 43168 43963: contig of 796 bp in length
* 43964 44063: gap of 100 bp
* 44064 44885: contig of 822 bp in length
* 44886 44985: gap of 100 bp
* 44986 45746: contig of 760 bp in length
* 45746 45845: gap of 100 bp
* 45846 46615: contig of 770 bp in length
* 46616 46715: gap of 100 bp
* 46716 47482: contig of 767 bp in length
* 47483 47582: gap of 100 bp
* 47583 48371: contig of 789 bp in length
* 48372 48471: gap of 100 bp
* 48472 49248: contig of 777 bp in length
* 49249 49348: gap of 100 bp
* 49349 50123: contig of 775 bp in length
* 50124 50223: gap of 100 bp
* 50224 51000: contig of 777 bp in length
* 51001 51100: gap of 100 bp
* 51101 51857: contig of 757 bp in length
* 51858 51957: gap of 100 bp
* 51958 52818: contig of 861 bp in length
* 52819 52918: gap of 100 bp
* 52919 53694: contig of 776 bp in length
* 53695 53794: gap of 100 bp
* 53795 54564: contig of 770 bp in length
* 54565 54664: gap of 100 bp
* 54665 55453: contig of 789 bp in length
* 55454 55553: gap of 100 bp
* 55554 56424: contig of 871 bp in length
* 56425 56524: gap of 100 bp
* 56525 57278: contig of 754 bp in length
* 57279 57378: gap of 100 bp
* 57379 58156: contig of 778 bp in length
* 58157 58256: gap of 100 bp
* 58257 59009: contig of 753 bp in length
* 59010 59109: gap of 100 bp
* 59110 59903: contig of 794 bp in length
* 59904 60003: gap of 100 bp
* 60004 60783: contig of 780 bp in length
* 60784 60883: gap of 100 bp

Query Match      8.3%; Score 51.6; DB 2; Length 63513;
Best Local Similarity 44.5%; Pred. No. 2.5;
Matches 186; Conservative 0; Mismatches 232; Indels 0; Gaps 0;

QY 164 TTGAACAATTTTCTACTATATTTTACAGATCATTTGGTTCAGGTTACACACAGT 223
Db 38964 TTGTTTANGTTTNTTANTTTTTTTATTTATTTAAATTTTTTTGGGGGAGTAA 39023

QY 224 CAGTATAGCTTATTTATTTATTTAGTTCCTCAAGAGTAGCTATTTATTTGACCAATAT 283
Db 39024 TATATATATTTATTTTATTTTATTTTAAAAAATAAATATTTTATTTTATTTT 39083

QY 284 TACTTTTATTCGCTGACGACGATTCATTCAGGTGACAGGTCATTTATGCA 343
Db 39084 TTTTATTTATTTTATTTATTTATTTATTTAAATTTATTTTATTTTATTTATTT 39143

QY 344 CTACAAAGGTTTTTAAATTTAAATGAGTAAATTTAATATTAATAAAGGTTGAAGGAA 403
```

```

Db      39144 TAAATATATTTTGTATATATATTTTAAATTTTATATTTTATTTTATTTTATTTA 39203
Qy      404 TGTTCCTTCGAACAACACTGTTGCAGCTTTACGATGTTAGTCAAGTAATGTTA 463
Db      39204 TATTTTATTTTATTTTAAATTTTAAATATATATTTAATTTATTTATTTATTTA 39263
Qy      464 TAGATATGTTCCATATATGATTAATAAGAGAGCTTTCAACCTTTTACCATATAT 523
Db      39264 TTAATATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTATTTATTTATTT 39323
Qy      524 ATGCAAGTCCTTCAGGTGCGAGATTAATGCTTAATACAGGAATAGTCTGTTT 581
Db      39324 AATTATATATTTTATTTATTTTATTTTATTTTATTTTATTTATTTATTTATTT 39381

RESULT 39
BX004827/c 119555 bp DNA linear PRI 11-APR-2003
LOCUS      Human DNA sequence from clone RP13-391G2 on chromosome X, complete
DEFINITION
ACCESSION  BX004827
VERSION     BX004827.18 GI:29824042
KEYWORDS   HTG.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE  1 (bases 1 to 119555)
AUTHORS   Ellwood, M.
TITLE      Direct Submission
JOURNAL    Submitted (11-APR-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: clones@wells.sanger.ac.uk
            humquerry@sanger.ac.uk; clone requests: clones@wells.sanger.ac.uk
            On Apr 11, 2003 this sequence version replaced gi:29603318.

COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMBEP; Information
on the WORMBEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormep This sequence
was generated from part of bacterial clone configs of human
Chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
RP13-391G2 is from the library RPCT-13.2 constructed by the group
of Pletier de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: PBACe3.6.
FEATURES
Source
1..119555
location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

```

BASE COUNT  30503 a 28003 c 27878 g 33171 t
ORIGIN
Query Match      8.3%; Score 51.6; DB 9; Length 119555;
Best Local Similarity 48.8%; Pred. No. 2.1;
Matches 168; Conservative 0; Mismatches 174; Indels 2; Gaps 1;

Qy      181 TATATTTTACAGATCATTTTCGGTTGAGGTTACAGACAGTCAGATATGCTTATTT 240
Db      86743 TATATGCTTATATATATTTTATATGATATATATATATATATATATTTATGAT 86684
Qy      241 ATTATGCTGTCAGAGATAGCTATTTATATGATGCCAATATACCTTTTATGCTGTT 300
Db      86683 ATATATGTTTATATATATATATTTTATATATATATATATATATATATATATG 86624
Qy      301 ACGCATGATTTTCATTACGTCGTCAGATTTGTCATATGACATACAAAGTTTAA 360
Db      86623 TATATATGTTTATATATATTTTATATGATATATATGTTTATATATATTTATAT 86564
Qy      361 TTAAATGAGATTAATTTATATATATATATATATATATATATATATATATATAT 418
Db      86563 TACATATGTTTGTATATATATATATATATATATATATATATATATATATAT 86504
Qy      419 AAACACTGTTGCACTTTTACGATGTTAGTCAAGTAATGTTATAGTATGTTCCAT 478
Db      86503 ATATATTTTATACATATTTTATATATATATATATATATATATATATATATAT 86444
Qy      479 ATATGATTAATAAGAGAGAGTTTCAAACTTTTACCATTTAT 522
Db      86443 TTAATGTTTATATATATATATATATATATATATATATATATATATATAT 86400

RESULT 40
CPU8070
LOCUS      CPU8070 10807 bp DNA linear BCT 07-MAY-1999
DEFINITION Chlamydomophila caviae cds1, cds2, copn and scs1 genes, complete cds.
ACCESSION  U08070
VERSION     U08070.1 GI:2444072
KEYWORDS
SOURCE      Chlamydomophila caviae
ORGANISM   Chlamydomophila caviae
REFERENCE  1 (bases 1 to 10807)
AUTHORS   Hsiao, R.-C., Pannekoek, Y., Ingerowski, E. and Bavoil, P.M.
TITLE      Type III secretion genes identify a putative virulence locus of
            Chlamydia
JOURNAL    Mol. Microbiol. 25 (2), 351-359 (1997)
MEDLINE    97426043
PUBMED     9282747
REFERENCE  2 (bases 1 to 10807)
AUTHORS   Hsiao, R.-C., Pannekoek, Y., Ingerowski, E. and Bavoil, P.M.
TITLE      Direct Submission
JOURNAL    Submitted (03-FEB-1997) Clinical Sciences, London School of Hygiene
            and Tropical Medicine, Keppel Street, London WC1E 7HT, UK
            On Sep 30, 1997 this sequence version replaced gi:2358255.

COMMENT
Location/Qualifiers
1..10807
/organism="Chlamydomophila caviae"
/mol_type="genomic DNA"
/strain="GPIC"
/specific_host="guinea pig"
/db_xref="taxon:83557"
/note="Submitted as Chlamydia psittaci
guinea pig inclusion conjunctivitis"
315..1397
/gene="cds1"
315..1397
/gene="cds1"
315..1397
/gene="cds1"
315..1397
/gene="cds1"
/function="cytoplasmic membrane component required for
type III (contact-dependent) secretion"

```



/note="similar to Spa40 of *Shigella flexneri*, Spa and  
Ssa of *Salmonella typhimurium*, and Yscu of *Yersinia*"  
/codon start=1  
/transl\_table=11  
/protein\_id="AA871508.1"  
/db\_xref="GI:2444073"  
/translation="MGEKTEKATPPKRLRDARKKQVAKSODPFAVTFIVSMPTTSEL  
ASFFKRLGSFLVSIPEKQAPINHDPKTLVYLVONCLVLTITSLPLGAAGFVGIIVG  
FLVVGPFSTEVKEDPKLPKPNIDNLKPKFVTLIELLSIKLFGALVILVTLAKN  
RVLPIETAGVSIVIGVILKQILYKAVTSIGIFELVAVALDVYQKNAKELKMEK  
FEVKEBETGNSDEIKGRROIAOEIAYEDTSOIKHASAVSNPKDIAVAGME  
KTAAPMIAKGINLRARIRIYEAKGVPIKRVNVLAHQIMDEKELKFTPESTYAI  
GELLYITSLNAQNPNNKNINQDNL"  
1397..3526  
/gene="cdc2"  
1397..3526  
/gene="cdc2"  
/function="cytoplasmic membrane component required for  
type III (contact-dependent) secretion"  
/note="similar to MxiA of *Shigella flexneri*, InvA and Ssa of  
*Salmonella typhimurium*, and LcrD of *Yersinia*"  
/codon start=1  
/transl\_table=11  
/protein\_id="AA871509.1"  
/db\_xref="GI:2358257"  
/translation="MNKLNFVSRTEFGGEAALNMINKSDLLIAMIIVMLIPL  
PPIVDIMITINLAVSVFLMALVYIPSAQLSVPELLITTMFRGINISSRLI  
LKAVAGHIVAFGDFVGVGVVVFIIITIIIOFTVMKGAERVAEVARPILDM  
PKGMADIDALPAGMIDAOQARDKRMIOKESHELGMQDMKFIKDVYAGVYISL  
NIVGLTIGVAMGMDLAQAHVYTLISDGLVSQIPSLISITAGITVTRSDKN  
TNLKEISQLVKPEPRLALLAGATLVGVFKGPVPSFSLIAVIFGLGVILLAKN  
NASKGASGATTVGAADAAGATAGDPDYALTLPVILEKDLSELIOQRKSGQS  
FIDMIPKMSQALYODIGIRYPGIHVTRDSPEGYMLINEVPYRGKVPNNVL  
TNVEENLKRYNLPTIYKNAAGPSAWSEDAKTLEKAIKYMPLFIITLISYF  
FHSSQELGICQVRSMIFEMERSPLVYEVTRILPLQGLTEIKRLVQESINDL  
RTLESSEWQTEKDYLVLTBYRSLKLIKISKSGQSAISVTLIDELIEMRG  
AIKQTSAGSYLADPDSVNLILKSMRTIPTPGGQPVLLTAIDVRRVRLIETE  
PDIIVISYOEIIPERIQPLGRIOIF"  
3557..4750  
/gene="copN"  
3557..4750  
/gene="copN"  
/function="protein substrate of type III secretory  
pathway, associated with outer membrane; putative  
regulatory function"  
/note="similar to MxiC of *Shigella flexneri*, InvE of  
*Salmonella typhimurium* and LcrE of *Yersinia*"  
/codon start=1  
/transl\_table=11  
/protein\_id="AA871510.1"  
/db\_xref="GI:2358258"  
/translation="MASAGAGLIGSQAIVDAVQAAMAKADQEVYASQESDISM  
IKDSQISNPOAATRTTKKEKROTLESRRKATQAPKESBDSADPADLTYTEN  
NAISQGLDSIRDSLDSSSEEDVDLVSKSDPQLQSVADLYVQTPPASKGALK  
DILIRAQNMQONQAVVGGKNILPASQEVASILNTSAPGLALYLEVTSDFNSCEQ  
LILTSQSYSEEMGVSSFILKMAADLSESGSIPAPKQVMTETRNQALVLY  
HFPEKTLPTLTKADGVTVPLDKPKVADTEPKLINDKPEIPASKKRGVRDIDGD  
TEAVTGMNLFPALRGTSPRLPASAEKROQLGTMMANALDAVINNEDYFKSTDPK  
PYFMS"  
4766..5206  
/gene="usc1"  
4766..5206  
/gene="usc1"  
/function="putative chaperone of type III secretory  
pathway"  
/note="similar to SycE of *Yersinia*"  
/codon start=1  
/transl\_table=11  
/protein\_id="AA871511.1"  
/db\_xref="GI:2444074"  
/translation="MQNQFROLESGLTKNTSLVPPKNAQCLIRSATVPOVLEED  
GNSGDIAGVITLGLPENVRERIFKALSVNASPQSNIGITLGGISQOLYISDLV  
NMSYLNQDKLFHYLNLFPMSHAKIWIALLTGNLPDLHVLGMYH"

CDS  
5263..6855  
/function="probable maltose polymer synthase"  
/note="orfE; similar to MalO (amylomaltase) of  
prokaryotes"  
/codon start=1  
/transl\_table=11  
/protein\_id="AA871512.1"  
/db\_xref="GI:2444075"  
/translation="MTPEFSKALRQIQNSPAKQSMKTLGIMPKHGI CLPLFSLRNSC  
GIEPDLIPMISMCRRGHGFOIIQIIPINDSGDSSPNSISVALNPILSLASLBH  
AQVAAANALTRPMOOLSKLPHYHYQVAKAEPLRDYQYVYKIGALDDEDEIIC  
EKETWLRPYTVFRSITKHLKAPVNNPRAVTDIKNFTEFEKQPODECSFSTLYQL  
CFQOMQVAFADNDNVFLKGDPLILISKSCDVWYRQFFSSGSGASAPDIYNTG  
ONMHLPIYNNHNLVQDNVTWVKARLYAENFSLYRLDHLVGLFRLVWMDTSGNGKFO  
PDDPKELPGQTDILQILASRMLPIGELDGSVPVDKETVLVGLIGCTRIPMBEN  
WEGDGNFPLGEYSPLSTVLSSTHSDSTLALMRHAKEXOKFAOPLGMRPTVLAE  
DQKHILTSKTSISIRHINIDYLYLCPULVSNNLKYERINMPGVTSKNWYRIKIP  
SVEIILTHDAFNANIDIFSKI"  
7313..9058  
/note="orfF"  
/codon start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AA871513.1"  
/db\_xref="GI:2444076"  
/translation="MAMSLRLVSLRSLQNDLIDIFAPVDPYLEPEIHRVYENOG  
PALFHNVOGASPEVLNLTGTOKRVQDIFSKYKGLIPQVILHSLSPQLSOLMKR  
NLLRGISGLRARRPLKEPKHKAFLNLSHQLMSPWDGGAFTLPLVYESSPS  
KITNLGHRNQRDRDRLTGLHPQIQKGGNHPEABEKNNLPLTVLNSNPILISA  
IAPLENISBLCLTFLQSKLHYKNDPYPHPLDYSEILILEGICNNRREPFG  
DHGGYVSLQDPEAFKRYKHKDAIYPTATYVQEDPYGNKQEYLSPLFVW  
MPEVROLKSGEAFALTAGVNERKESLTSRIILEGQSLTKPLMITDHDV  
LDNPFLETLTIRIYBERDLIFESSTNDLTDTYGPKLKSGKALFMGIGPARIDP  
HKRGSLEPMTITMGPCBGLVLETLQOVNIDALINHPDLSMPLVLTENNELT  
ASSKDFUKFTFLAIPATDIAVAFNSVAHRPVTPIILNSLKKHPYKVEADETT  
IQVRSHNBYPKFY"  
9164..10219  
/note="orfG"  
/codon start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AA871514.1"  
/db\_xref="GI:2444077"  
/translation="MKKTKSKLILTLISLIFGLLTQSAQDPDTFOTFASQEPVI  
YSKQGGDDTLKVLCDALISAQTSIPRIYVLSPEITSLANQARARNVAIHKXIA  
KITEPRKNRVTYVHEPABEKRLMKOKALAIIDKRNALGSANTHTSFQDSSNITGL  
KSKELCRYIDESCQGVHGNQAYFSLPDGQKALSAVLRTLRTAKTIRLMMFL  
TYPVDFRELNEAKRGVDVKILIDKDFKLSIKQIOALKDSNLSLHRTTTRYRLHNF  
AVVDOKILINGSVWSEAGFCINSEBMLIIDLNTDKQLKLDKIMDKLEKQSSILYEV  
SDDEKVIOLPEKCA"

BASE COUNT 3266 a 2239 c 1998 g 3304 t  
ORIGIN  
Query Match 8.3%; Score 51.4; DB 1; Length 10807;  
Best Local Similarity 47.6%; Pred. No. 4.4;  
Matches 185; Conservative 0; Mismatches 201; Indels 3; Gaps 1;  
QY 19 AAAACAGAAACCAACCCCGAAACGACGACAGAACTGTTCTGAAGGAGTGCCT 78  
DB 324 AAAACAGAAAGCGAGCCCGAAGCGCTTAGAGACCCGAGAAAAAAGCCAGTAGCA 383  
QY 79 AAATCAGAAAGGTTACTTAAGCACTTACTACTGACAGACGAGATCGTGGGCTGTATT 138  
DB 384 AAATCTCAAGATTTTCTTCCCGCGTTACCTTTATGCTTTCGATGTTACAGACATTTTCT 443  
QY 139 TATTCAGCGTAATGGACGTCATTTTGAACAATTTTCTACTATATTTTATACGAATCA 198  
DB 444 CTAGCATCGTTTTCGCAAAACATCTGTAGTTTCTACTATTTCTATTAAAGCAAGCC 503  
QY 199 TTTCGGTTTGAAGTTACAGACAGTCAGTATATGCTTATTTATTTATGTTGCTCAAGAG 258  
DB 504 CTTATTAATCATGACCTTAAGTCACGCTATACTATTACAAAATGCTTAGTTTATTT 563

```

Qy 259 ATAGCTATTTTATGATGCAATATTAATTTTATGCTGTACGGCATGGATTTCATTA 318
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 564 CTAACTACTTCCCTT---CCCTACTGTGTGCTGTGGATTTGTGGTATTATTTGCGA 620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 319 CGTGACAGTTGGTGCAATTATGACTACAAAGTTTAAATTAAATGAGTAATTT 378
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 621 TTCTGTGTGTAGGCCCACTTTTCTACAGAGGTTTAAAGCTTAATAAAGTTC 680
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 379 AATAAATAAAGGTTGAAGGAATGTT 407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 681 AATCTATTGCACTTGAACAAAATTT 709
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: October 4, 2003, 04:39:26  
 Job time : 2061 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2003, 18:14:56 ; Search time 217 Seconds  
(without alignments)  
7737.560 Million cell updates/sec

Title: US-10-010-160-1

Perfect score: 622

Sequence: 1 atgcgtgatccagtaa.....aattatgcagtcgcagatc 622

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_190un03.\*

- 1: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
- 2: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
- 3: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
- 4: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
- 5: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
- 6: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
- 7: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
- 8: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
- 9: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
- 10: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
- 11: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
- 12: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
- 13: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
- 14: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
- 15: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
- 16: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
- 17: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
- 18: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
- 19: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
- 20: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
- 21: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
- 22: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
- 23: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
- 24: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*
- 25: /SIDSI1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	622	100.0	622	24	ABK97884
2	54.8	8.8	8056	25	ABZ10100
3	54.8	8.8	9052	24	ABZ12145
4	54.4	8.7	8056	25	ABZ10246
5	53.6	8.6	1501	25	ABZ10188
6	53	8.5	945	19	AAK14127
7	53	8.5	1024	24	ABK6125
8	53	8.5	1086	18	AAK67941

9	52.8	8.5	7167	24	ABL32401	Human immune syste
10	52	8.4	5297	22	AAK46514	Tumour suppressor
11	52	8.4	7128	24	ABL33559	Human immune syste
12	52	8.4	40324	24	ABO67150	Human angiogenesis
13	51.8	8.3	5523	24	ABL32164	Human immune syste
14	51.8	8.3	73334	24	ABL92318	Chemically treated
15	51.8	8.3	73334	24	ABL34124	Human immune syste
16	51.2	8.2	12025	24	ABL33299	Human immune syste
17	51	8.2	6118	24	ABL33031	Human immune syste
18	51	8.2	7921	24	ABL33971	Human immune syste
19	50.6	8.1	6126	22	AAK46574	Tumour suppressor
20	50.6	8.1	6126	24	ABL33831	Human immune syste
21	50.6	8.1	11047	22	AAK45480	Chemically pretrea
22	50.6	8.1	11047	24	ABL33985	Human immune syste
23	50.6	8.1	11047	24	ABK28402	DNA transcription
24	50.2	8.1	9118	24	ABO66989	Human angiogenesis
25	50.2	8.1	9543	24	ABN80048	Human chemically m
26	50.2	8.1	513445	22	AAK161373	Soybean 318013 reg
27	49.8	8.0	5935	22	AAK45426	Chemically pretrea
28	49.8	8.0	40862	24	ABL34072	Human immune syste
29	49.8	8.0	1038602	20	AAK01425	Complete genome se
30	49.2	7.9	61020	22	AAK46788	Tumour suppressor
31	48.8	7.8	5542	24	ABL34021	Human immune syste
32	48.8	7.8	5857	24	AAK56347	Chemically pretrea
33	48.6	7.8	18855	24	ABL32610	Human immune syste
34	48.4	7.8	11787	24	ABL92243	Chemically treated
35	48.4	7.8	19345	24	ABN80017	Human chemically m
36	48.2	7.7	12138	24	ABK40034	Human immune syste
37	48.2	7.7	12138	24	ABL33629	Human immune syste
38	48.2	7.7	12590	24	ABL33920	Human immune syste
39	48.2	7.7	18512	24	ABL32977	Human immune syste
40	48.2	7.7	82952	24	ABN85766	Arabidopsis yellow
41	48	7.7	13427	24	ABL33937	Human immune syste
42	47.8	7.7	524	24	ABO34338	Oligonucleotide fo
43	47.8	7.7	524	24	ABO34339	Oligonucleotide fo
44	47.8	7.7	5126	24	ABL70493	Chemically treated
45	47.8	7.7	6310	24	ABL70321	Chemically treated

## ALIGNMENTS

## RESULT 1

ABK97884 ID ABK97884 standard; DNA; 622 BP.

XX ABK97884;

XX 10-OCT-2002 (first entry)

XX DNA encoding Lawsonia intracellularis antigenic flhB protein.

XX Immunogen; flhB; flhR; ntrC; glhH; motB; clyC; ytfW; ytfH;

XX porcine; pig; avian; bird; porcine proliferative enteropathy; PPE;

XX intestinal adenomatosis complex; porcine intestinal adenomatosis;

XX PIA; necrotic enteritis; proliferative haemorrhagic enteropathy;

XX regional ileitis; haemorrhagic bowel syndrome; vaccine; antibacterial;

XX porcine proliferative enteritis; Campylobacter spp.-induced enteritis;

XX gene; ds.

XX Lawsonia intracellularis.

XX WO200238594-A1.

XX 16-MAY-2002.

XX 09-NOV-2001; 2001WO-AU01462.

XX 10-NOV-2000; 2000AU-0001381.

XX 17-NOV-2000; 2000US-249596P.

XX (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.

XX (AUPO-) AUSTRALIAN PORK LTD.



CC subclasses, diagnosis, prognosis, treatment and/or monitoring of  
CC hematopoietic cell proliferative disorders. The present method enables  
CC a highly specific classification of hematopoietic cell proliferative  
CC disorders allowing for improved and informed treatment of patients.

XX Sequence 8056 BP; 3711 A; 371 C; 371 G; 3603 T; 0 other;

Query Match 8.8%; Score 54.8; DB 25; Length 8056;  
Best Local Similarity 47.2%; Pred. No. 0.0035;  
Matches 167; Conservative 0; Mismatches 187; Indels 0; Gaps 0;

QY 160 CATTGGAACAATTTCTACTATATTTTACAGAAATCTTCGGTTGAGGTACGCA 219  
DB 6351 CGTATATAATAATTTTATTTATTTATTTAAAAAAATATATATATATATATAT 6410  
QY 220 CAGTCAGTATATGCTTATTTATTTATTTGCTCAAGATAGCTATTTATGATGCA 279  
DB 6411 ATATATATTTATTTAATTTAATTTATTTATTTAATTTAATTTTATTTATTTAAA 6470  
QY 280 ATATTTACTTTTATTTGCTGTTACGAGATTCATTTACGTTACAAAGTTGTCATTA 339  
DB 6471 ATATATAATATCGTTTTTTTATTTGTTATTTAATTTTCGTAATTTTCGTAAGAA 6530  
QY 340 TGCAGCTCAAAAGGTTTTTAAATTTAAATGAGTAATTTAATATATATTTAAAGGTTGAAA 399  
DB 6531 AACGTTATTTTAAATTTTAAATATATTTATTTATTTAAATTTAAATTTAAAGAAAA 6590  
QY 400 GGAATGTTGCTGCTCAACAACACTGTTGCACTTTTACGTAAGTTAGTTCAGTAAT 459  
DB 6591 AATATTTATTTATTTTATTTTATTTATTTATTTATTTATTTATTTTATTTTACGTTTAA 6650  
QY 460 GTATATAGTATTTGTCATATATGATTTAATAAGAGAGTTTCAAACTTTTAA 513  
DB 6651 AATTTTAAATTTTATTTAATTTTAAATTTTAAAAAAATTTTATTTTACGTTTAA 6704

RESULT 3  
ABLJ2145  
ID ABLJ2145 standard; DNA; 9052 BP.

XX ABLJ2145;  
XX  
XX 26-MAR-2002 (first entry)  
XX  
XX Human immune system associated gene SEQ ID NO: 118.

XX Human; immune system disease; cytosine methylation; antiasthmatic;  
XX antiarteriosclerotic; anti-anemic; cytosolic; nootropic;  
XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
XX antineumatic; antiarthritic; antidiabetic; antiporotic;  
XX antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
XX acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
XX neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
XX gene; ds.

OS Homo sapiens.

XX WO200200928-A2.

XX 03-JAN-2002.

XX 02-JUL-2001; 2001WO-EP07537.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful

PT for diagnosis and treatment of diseases associated with abnormal  
PT cytosine methylation  
XX  
XX Claim 1; SEQ ID NO 118; 32pp + Sequence Listing; German.

CC The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention.

XX Sequence 9052 BP; 2166 A; 112 C; 2104 G; 4670 T; 0 other;

Query Match 8.8%; Score 54.8; DB 24; Length 9052;  
Best Local Similarity 47.2%; Pred. No. 0.0036;  
Matches 167; Conservative 0; Mismatches 187; Indels 0; Gaps 0;

QY 138 TTATTCAGGCGTATGGAGCGTCATTTTGAACAATTTTCTACTATATTTTACAGATC 197  
DB 799 TTTTTCAGATGAAAAAGAGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 858  
QY 198 ATTCGTTTGAAGTTTACAGACAGTCAGTATATGCTTATTTATTTATTTATTTGCTCAGA 257  
DB 859 TTATTTTATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 918  
QY 258 GATAGCTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 317  
DB 919 TTGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 978  
QY 318 ACGTTCAGATTTGCTGTCATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 377  
DB 979 ATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1038  
QY 378 TAAATATATAAAAGGTTGAAGAGATTTGCTTCAACAAACACTGTTGACATTTT 437  
DB 1039 TGTGAATAATADAGATTTTAAATGATGATATTTGTAAGTATTTTATTTTATTTTATTTGATG 1098  
QY 438 ACGTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 491  
DB 1099 AGGTAAAGTATTTTATTTTGAATTTTGTGTTTAAATTTTAAATTTGTTGATATGA 1152

RESULT 4

ABZ10246  
ID ABZ10246 standard; DNA; 8056 BP.

XX ABZ10246;  
XX

XX 16-JAN-2003 (first entry)  
XX

XX Haematopoietic cell proliferation disorder related DNA sequence #386.

XX Human; haematopoietic cell proliferation disorder; cytosolic;  
XX gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;  
XX cytosine methylation state; gene; ds.

OS Homo sapiens.

XX WO20027272-A2.

XX 03-OCT-2002.

XX 26-MAR-2002; 2002WO-EP03401.

XX 26-MAR-2001; 2001US-278333P.

XX (EPIG-) EPIGENOMICS AG.

XX Berlin K, Braun A, Distler J, Guefig D, Howe A, Mueller J;

XX Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;

PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T;  
 PI Pelet C, Schwöpe I, Ziebarth H;  
 DR WPI; 2003-018942/01.  
 XX  
 PT Detecting and differentiating between hematopoietic cell proliferative  
 PT disorders, comprises contacting a target nucleic acid with a reagent  
 PT that distinguishes between methylated and non-methylated CpG  
 PT dinucleotides -  
 PS  
 PS Claim 28; SEQ ID 386; 117pp; English.  
 XX  
 CC The present invention describes a method for detecting and  
 CC differentiating between haematopoietic cell proliferative disorders  
 CC associated with at least 1 gene and/or their regulatory regions in a  
 CC subject. The method comprises contacting a target nucleic acid in a  
 CC biological sample obtained from the subject with at least 1 reagent,  
 CC which distinguishes between methylated and non-methylated CpG  
 CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ1118  
 CC represent specifically claimed nucleotide sequences from the present  
 CC invention. Oligonucleotides from the present invention can be used: for  
 CC differentiating between healthy haematopoietic cells and proliferative  
 CC disorder haematopoietic cells; for differentiating between acute  
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for  
 CC determining the cytosine methylation state and/or single nucleotide  
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder  
 CC related sequences and their complements; and as primers for the  
 CC amplification of haematopoietic cell proliferation disorder related  
 CC DNA sequences. The nucleotide sequences from the present invention can  
 CC also be used for detecting a predisposition to, differentiation between  
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of  
 CC haematopoietic cell proliferative disorders. The present method enables  
 CC a highly specific classification of haematopoietic cell proliferative  
 CC disorders allowing for improved and informed treatment of patients.  
 CC  
 SQ Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 other;  
 Query Match 8.7%; Score 54.4; DB 25; Length 8056;  
 Best Local Similarity 46.7%; Pred. No. 0.0043;  
 Matches 172; Conservative 0; Mismatches 196; Indels 0; Gaps 0;  
 QY 162 TTTTGAACAATTTCTACTATATTTTACAGAAATCTTGGTTGAGTTACAGACACA 221  
 Db 6553 TTAATTAATTAATTTTATTAATTTATTAATTAATTAATTAATTAATTAATTAAT 6412  
 QY 222 GTGAGTATATCTTATTTATTTATTTGTTGCTCAAGAGATAGCTATTTTATGATGCCAAT 281  
 Db 6413 ATAAATTAATTAATTAATTTTATTTTAAATTAATTTTATTTATTTATTAATTAAT 6472  
 QY 282 ATTAATTTTATTTGCTGTTAGGCGATGATTTTCATTACGTCAGCAAGTTGGCTATATG 341  
 Db 6473 AATAAATAATTTTATTTTATTTGTTTATTTTAAATTTTGGAAATTTTGTATGAAAAA 6532  
 QY 342 GACTACAAAGGTTTAAATTTAAATGAGTAATTTTAAATTAATTAATTAATTAATTAATTAAT 401  
 Db 6533 TGTATTTATTAATTTTAAATTAATTTATTTATTTTAAATTAATTAATTAATTAATTAATTAAT 6592  
 QY 402 AATGTTTGTCTTCAACAACAAGCTTTGACCTTTTACGTAAGTTAGTTCAAGTAATTTGT 461  
 Db 6593 ATTAATTTTATTTTATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 6652  
 QY 462 TATAGATATTTGTCATATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 521  
 Db 6653 TTTTAAATTTTATTAATTAATTTTAAATTAATTAATTTTATTTATTTATTTATTTATTTATTT 6712  
 QY 522 TTATGCAA 529  
 Db 6713 ATTATATA 6720

RESULT 5  
 ABZ10188  
 ID ABZ10188 standard; DNA; 1501 BP.

XX  
 AC ABZ10188;  
 XX  
 DT 16-JAN-2003 (first entry)  
 XX  
 DE Haematopoietic cell proliferation disorder related DNA sequence #328.  
 XX  
 KW Human; haematopoietic cell proliferation disorder; cytostatic;  
 KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;  
 KW cytosine methylation state; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 EN WO200277272-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PE 26-MAR-2002; 2002WO-EP03401.  
 XX  
 PR 26-MAR-2001; 2001US-278333P.  
 PA (EPIG-) EPIGENOMICS AG.  
 XX  
 PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;  
 PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;  
 PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T;  
 PI Pelet C, Schwöpe I, Ziebarth H;  
 DR WPI; 2003-018942/01.  
 XX  
 PT Detecting and differentiating between hematopoietic cell proliferative  
 PT disorders, comprises contacting a target nucleic acid with a reagent  
 PT that distinguishes between methylated and non-methylated CpG  
 PT dinucleotides -  
 PS  
 PS Claim 28; SEQ ID 328; 117pp; English.  
 XX  
 CC The present invention describes a method for detecting and  
 CC differentiating between haematopoietic cell proliferative disorders  
 CC associated with at least 1 gene and/or their regulatory regions in a  
 CC subject. The method comprises contacting a target nucleic acid in a  
 CC biological sample obtained from the subject with at least 1 reagent,  
 CC which distinguishes between methylated and non-methylated CpG  
 CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ1118  
 CC represent specifically claimed nucleotide sequences from the present  
 CC invention. Oligonucleotides from the present invention can be used: for  
 CC differentiating between healthy haematopoietic cells and proliferative  
 CC disorder haematopoietic cells; for differentiating between acute  
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for  
 CC determining the cytosine methylation state and/or single nucleotide  
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder  
 CC related sequences and their complements; and as primers for the  
 CC amplification of haematopoietic cell proliferation disorder related  
 CC DNA sequences. The nucleotide sequences from the present invention can  
 CC also be used for detecting a predisposition to, differentiation between  
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of  
 CC haematopoietic cell proliferative disorders. The present method enables  
 CC a highly specific classification of haematopoietic cell proliferative  
 CC disorders allowing for improved and informed treatment of patients.  
 CC  
 SQ Sequence 1501 BP; 604 A; 0 C; 130 G; 767 T; 0 other;  
 Query Match 8.6%; Score 53.6; DB 25; Length 1501;  
 Best Local Similarity 47.7%; Pred. No. 0.0047;  
 Matches 187; Conservative 0; Mismatches 204; Indels 1; Gaps 1;  
 QY 131 TTGCTATTTATTTGCGGTATGCGATGATTTTGAACAATTTCTACTATATTTTAA 190  
 Db 669 TTTTAATTAATTTATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 728  
 QY 191 CAGATCATTTGCGTTTGAAGTTACAGACAGTCAGTATATGCTTTATTTATTTATTTATTTATTT 250  
 Db 729 TTGTTTATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 788

QY 251 CTCAAGAGTAGCTATTATTATGATGCCAATATCTTTTATGCTGTACGGCAGGA 310  
 Db 789 AATTAATAATATTTTGTGTTTATGTTATGTTATTTTGTGATGTTATTTGTTAA 848  
 QY 311 TTTTATTACGTCGACAAAGTGTGCTATTGACTACAAAGCTTTTAAATTAATGA 370  
 Db 849 TTTTAAATATGTTTAAATGTTGTAATAAATAAATAAATAAATGTTATTAATGATT 908  
 QY 371 GTAATTAATATATAAAGGTTGAAGAAGTGTGCTTCAACAAACACTGTTC 430  
 Db 909 TTTTATTTTAAAT-TAAATTAATTTGATGTTGTTAAATTTGTTGTTT 967  
 QY 431 GACTTTTACGTAAGTGTGCTCAAGTAATGTTATAGTAATGTTCAATATGATTAA 490  
 Db 968 TAAATTTAATATTTAAATTTAAATATTAATGAAATGAAATTTGAAATTTAAATTTG 1027  
 QY 491 AAGGAGAGTTTCAAACTTTTACCAATTAT 522  
 Db 1028 AAATGTAATTTTAAATTTTAAATTTTATTT 1059

## RESULT 6

AAK14127/c  
 ID AAK14127 standard; DNA; 945 BP.

AAK14127;

31-MAR-1999 (first entry)

H. pylori GHPO 243 gene.

GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;

peptic ulcer disease; ss.

Helicobacter pylori.

Key Location/Qualifiers

FT CDS 61..640

FT /\*\*tag= a

W09843478-A1.

08-OCT-1998.

01-APR-1998; 98MO-US06371.

29-JUL-1997; 97US-0902615.

01-APR-1997; 97US-0833457.

24-JUN-1997; 97US-0881227.

(HUMA-) HUMAN GENOME SCI INC.

(INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.

Al-Garawi A, Kleantous H, Miller C, Oomen RP, Tomb J;

WPI; 1998-542293/46.

P-PSDB; AAM98408.

New isolated Helicobacter polynucleotides - used to develop products for the diagnosis, prevention and treatment of Helicobacter infections and gastrointestinal diseases

PS Claim 1; Page 679-680; 2054pp; English.

This sequence represents a polynucleotide of the invention. It was isolated from Helicobacter pylori and encodes a H.pylori GHPO protein. The polypeptides can be used for preventing or treating Helicobacter infections, and gastroduodenal diseases associated with these infections, including acute, chronic, and atrophic gastritis, and peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be used for the production of antibodies. The products can also be used for detection and diagnosis.

XX SQ Sequence 945 BP; 298 A; 186 C; 169 G; 292 T; 0 other;

Query Match 8.5%; Score 53; DB 19; Length 945;

Best Local Similarity 46.9%; Pred. No. 0.0059;

Matches 201; Conservative 0; Mismatches 225; Indels 3; Gaps 1;

QY 19 AAAACAGAGAAAGCAACCCGGAACGACGACAGAGCTGTTCTGAAGGAGTGCCT 78  
 Db 845 AAAACGAACTCCCTAGCGGAAATAATCCAAAAGCCAGAGAAAGGCAATGTGCT 786  
 QY 79 AAATCAGAAAGGTTACTAAAGCAATGACTACTGACAGAGAGTCTGGGCTGTAAT 138  
 Db 785 AAGAGCATGAGAGTGGGGGTTTGGGGTTATGGCCGGCTAATAGATTTTGT 726  
 QY 139 TATTCAGGCGTAATGGAGCTCATTTTGAACAATTTTCTACTATATTTTACGAATCA 198  
 Db 725 TTTTATATATATGATGGGATGAGCTTTAGCGAAATGATGCGCATGTGTGAAGA--T 669  
 QY 199 TTTGCGTTGAGGTTACAGCAGAGTCAATATGCTTATTTATTTATGTTGCTCAGAG 258  
 Db 668 TTTTCCCTAGATTTCACTAAGAAAGGCTTCAAGAGCTGTTTAACTAGCTTAAAGAC 609  
 QY 259 ATAGCTATTTTATGATGCAATATTAATTTATGCTGTACGGCAGATTTCAATTA 318  
 Db 608 ACTTTTATTTGCTTTTACCGATTTTATCATTTTATGATGTTGGGCTTTTATCTAAT 549  
 QY 319 CGGTACAGTGTGTGCTATTTAGCTACCAAGCTTTTAAATTAATGAGTAATTT 378  
 Db 548 GTCTTGCAATTTGGCTGGCTCTTGGCCCTTAAAGTCAATGAGCTTAAATTTTCTAAATC 489  
 QY 379 AATATAATAAAGGGTTGAAGAAGATGTTGCTTCCAAACAACTGTGCACTTTTA 438  
 Db 488 AACCCATCATATGCGGTCAAAAACCTTTTCTTTTAAAAAGCTCTTGATGGAGTTTG 429  
 QY 439 CGTAGTTTA 447  
 Db 428 ATCACCCTTA 420

## RESULT 7

ABX6125  
 ID ABX6125 standard; DNA; 1024 BP.

ABX6125;

07-MAY-2003 (first entry)

Helicobacter pylori selected interacting domain (SID) DNA #724.

Protein-protein interaction; ulcer; selected interacting domain;

KM SID; gene; ds.

Helicobacter pylori.

W020026501-A2.

29-AUG-2002.

28-DEC-2001; 2001MO-BP15428.

02-JAN-2001; 2001US-259302P.

(HYBR-) HYBRIGENICS.

(INSP) INST PASTEUR.

Legrain P, Rain J, Colland F, De Reuse H, Labigne A;

WPI; 2002-674910/72.

P-PSDB; ABUS1381.

New complexes of protein-protein interactions in Helicobacter pylori, useful for identifying modulating compounds for treating or preventing

PT ulcers in mammals -

XX Claim 7, Page 261; 642bp; English.

CC The invention describes a complex of protein-protein interactions in  
 CC Helicobacter pylori selected from 421 complexes given in the  
 CC specification. The complex of protein-protein interactions are useful  
 CC for screening for agents which modulate the interaction of proteins.  
 CC Modulating compounds which binds to a targeted bacterial protein may be  
 CC used for treating or preventing ulcers in a human or animal. This  
 CC sequence encodes a selected interacting domain (SID), identified via  
 CC protein-protein interactions.

XX Sequence 1024 BP; 312 A; 187 C; 203 G; 322 T; 0 other;

Query Match 8.5%; Score 53; DB 24; Length 1024;  
 Best Local Similarity 46.9%; Pred. No. 0.006;  
 Matches 201; Conservative 0; Mismatches 225; Indels 3; Gaps 1;

19 AAAACAGAGAAAGCAACCCGAAACGACGTCAGAGAGCTGCTTGAAGGAGGTCCCT 78  
 141 AAAACCGAATCTCCCTAGCGCGAATAAATCCAAAAGCCAGAGAAAGGCAATGTGCT 200  
 79 AAATCAGAGAGGTACTTAAGCATTAAGCTAGCAGAGAGGATGCTGGGCTTGTATT 138  
 201 AAGACATGAGAGTGTGGGGGTTTGGGGTTATGGCCGGCTAATTAATTTTGT 260  
 139 TATTCAGGCGTAAATGGGACGTCATTTTGAACAATTTCTACTATATTTTACAGATCA 198  
 261 TTTTATATATGTTGGTGGATGCTTGAAGAAATGTAATCGCATGTGTGAAGA--T 317  
 159 TTTGGTTGAGGTACAGCAGCATATATGCTTATATTTATTTGTTGCTCAAG 258  
 318 TTTTCCCTAGATTAGTAAGAAAGCGTCAAGAGCTGTTTAACCACTGCTAAAGC 377  
 259 ATAGCATTTTATTTATGATGCAATATTAATTTATTTGCTGTACGCAATGATTCATTA 318  
 378 ACTTTTATTTGCTTTTACCGATTTTAAATCATTTTATAGTGTGTGCGCTTTATCTAAT 437  
 319 CGTGACAAAGTGTGTCATTTATGACATCAAAAGGTTTAAATTAATGAGTAATTT 378  
 438 GTCTTGAATTTGGTGGCTCTTTTCCCTTAAAGCATTTGACCTTAAATTTCTTAAATC 497  
 379 AATATATATAAAGGTTGAAAGAAAGTTTGGTCTTCAACAACAAGCTTGTGCACTTTTA 438  
 498 AACCTATCATATGGGTCAAAAACCTTTTCTTTTAAAAAAGCTCTTGATGGAGTTTG 557  
 439 CGTAGTTTA 447  
 558 ATCACCCTTA 566

RESULT 8

AAT67941 ID AAT67941 standard; DNA; 1086 BP.

XX AAT67941;

DT 14-JUN-1997 (first entry)

XX H. pylori flagella-associated protein ORF 04gell130rfs.

XX Flagellum; vaccine; prevention; treatment; infection; identification;  
 KW binding compound; bacterium; life cycle; activator; bacterial inhibitor;  
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope; ds.

XX Helicobacter pylori.

XX Key Location/Qualifiers

XX CDS 1..1086

FT /\*tag= a

FT /note= "no stop codon given"

PM W09640893-A1.

XX 19-DEC-1996.

XX 06-JUN-1996; 96WO-US09122.

XX 01-APR-1996; 96US-0830405.

XX 07-JUN-1995; 95US-0487032.

XX (ASTR ) ASTRA AB.

PI Berglindh OT, Smith D, Mellgaard BL;

XX WPI; 1997-052306/05.

DR P-PSDB; AAM20688.

PT Helicobacter pylori nucleic acid sequences and related

XX polypeptide(s) - useful for vaccines to treat or prevent H. pylori

XX infection, and to detect Helicobacter

XX Claim 1; Pages 801-802; 1481bp; English.

XX The present sequence encodes a Helicobacter pylori flagella-

XX associated protein.

XX The protein may be used in a vaccine to prevent or treat H. pylori

XX infection or to identify H. pylori polypeptide binding compounds,

XX useful as potential H. pylori life cycle activators or inhibitors.

XX The genomic sequence of H. pylori (ATCC 55679) was determined from

XX overlapping contigs generated by mechanically shearing the bacterial

XX DNA. The sequences were analysed for ORF of at least 180 nucleotides,

XX and the predicted coding regions defined by computer evaluation. To

XX identify likely H. pylori antigens for vaccine development, the amino

XX acid sequences predicted from various ORF were analysed for significant

XX homology to other known or exported membrane proteins. Having identified

XX and determined the sequences of interest, particular regions can be

XX isolated from H. pylori by PCR amplification for recombinant polypeptide

XX production, e.g. in E. coli hosts.

XX Sequence 1086 BP; 335 A; 193 C; 229 G; 329 T; 0 other;

Query Match 8.5%; Score 53; DB 18; Length 1086;  
 Best Local Similarity 46.9%; Pred. No. 0.006;  
 Matches 201; Conservative 0; Mismatches 225; Indels 3; Gaps 1;

19 AAAACAGAGAAAGCAACCCGAAACGACGTCAGAGAGCTGCTTGAAGGAGGTCCCT 78  
 28 AAGACCGAATCTCCCTAGCGCGAATAAATCCAAAAGCCAGAGAAAGGCAATGTGCT 87  
 79 AAATCAGAGAGGTACTTAAGCATTAAGCTAGCAGAGAGGATGCTGGGCTTGTATT 138  
 88 AAGACATGAGAGTGTGGGGTTTATAGGTTATGCTGGCTAATGAGATTTTGT 147  
 139 TATTCAGGCGTAAATGGGACGTCATTTTGAACAATTTCTACTATATTTTACAGATCA 198  
 148 TTTTATATATGTTGGTGGATGCTTTAGCGAGATGATCGCATGTGTGAAGATTTTC 207  
 199 TTTGGGTTGAGGTTACAGACAGTCAATATTCCTTATTTATTTATTTGTTGCTCAAG 258  
 208 TCCC--TAGATTTTACCAAGAAAGGCTTCAAGAGCTTTTAAACAGCTGGCTTAAGAC 264  
 259 ATAGCTATTTATTTATGAGCAATATTAATTTTATTTTATGCTGTTAGGATGATTTCA 318  
 265 ACTTTTATTTATTTGCTTTTGGCTTTTATATCATTTTATATGTTGGGCTTTTGTCTAAT 324  
 319 CGTGACAAAGTGTGATTAATGACATCAAAAGGTTTAAATTTAAATGAGTAATTT 378  
 325 GTCTTGAATTTGGCTGGCTTTTCCCTTAAAGTCAATGAGCTTAAATTTCTTAAATC 384  
 379 AATATATATAAAGGTTGAAAGAAAGATGTTGCTTCAACAACAAGCTTGTGCACTTTTA 438  
 385 AACCTATCATATGGGTCAAAAACCTTTTCTTTTAAAAAAGATCCTTGATGGAGTTTG 444  
 439 CGTAGTTTA 447



Db 445 ATCACTTTA 453

# RESULT 9

ABL32401  
ID ABL32401 standard; DNA; 7167 BP.

AC ABL32401;

DT 26-MAR-2002 (first entry)

XX Human immune system associated gene SEQ ID NO: 374.

XX Human, immune system disease; cytosine methylation; antiasthmatic;  
XX antiarteriosclerotic; antianemic; cytosolic; noctropic;  
XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
XX antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
XX antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
XX acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
XX neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
XX gene; ds.

OS Homo sapiens.

XX WO200200928-A2.

PD 03-JAN-2002.

XX 02-JUL-2001; 2001WO-EP07537.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPISNOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful  
PT for diagnosis and treatment of diseases associated with abnormal  
PT cytosine methylation

PS Claim 1; SEQ ID NO 374; 32bp + Sequence Listing; German.

XX The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC disease. The present sequence is a gene of the invention.

XX Sequence 7167 BP; 2279 A; 47 C; 1050 G; 3791 T; 0 other;

XX Query March 8.5%; Score 52.8; DB 24; Length 7167;

XX Best Local Similarity 46.3%; Pred. No. 0.01; 202; Indels 0; Gaps 0;

XX Matches 174; Conservative 0; Mismatches 202; Indels 0; Gaps 0;

QY 234 TTTATTTATTTATGCTCTCAAGATGCTATTTTATGTCGCAATTTACTTTTAT 293

Db 709 TATATTTGTTAGTTTATGAGTTTAAAGTATGATTTATATTTTATTTTATG 768

QY 294 TCGCTTAAAGGCAATTCATTACGTGTAACAAGTGGTCATTATGACAAAGT 353

Db 769 AGATTTTATGCTTATGATTTTATTTTATTTTATTTTATTTTATTTTAAAT 828

QY 354 TTTTAAATTTAATGAGTAATTTATTTATTTAATTTAATTTAATTTTATTTTATTTT 413

Db 829 ATTATGTTTATGAGTAATTTATTTTATTTTAAATTTAAGGTTTGAATGTTTATTTATTTT 888

QY 414 TCACAAACACTTGTTCAGCTTTTACGTTTATGTTCAAGTATTTGTTATGATTTGT 473

Db 889 TGCTTTATATATGTTTATATGTTTATGAGTTTTCGATGATTTGCTTAAGTTTATA 948

QY 474 TCCATATATGATTTATTAAGAGAGAGGTTTCAAACTTTTACCATTTATATGCAATCC 533

Db 949 ATATTTTCGGGTTTTTATGAGATTTTAAAGTTTATTTTAAATTTTAAAGTTTTTATGATTA 1008

QY 534 TTCAGGTGCGCAGATTTATGCTTAAATACGAAATGCTCTTTTATATACGTTAT 593

Db 1009 TTTTGAATTTTGAATAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATG 1068

QY 594 TCCATGACCAATTTAT 609

Db 1069 TTTTATGAGAGTTT 1084

# RESULT 10.

AAS46514  
ID AAS46514 standard; DNA; 5297 BP.

AC AAS46514;

DT 18-DEC-2001 (first entry)

XX Tumour suppressor gene derived chemically modified sequence #236.

XX Human, tumour suppressor gene; oncogene; antitumour; cytostatic;  
XX cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;  
XX cytosine methylation; ds.

OS Homo sapiens.

XX WO200168912-A2.

XX 20-SEP-2001.

XX 15-MAR-2001; 2001WO-EP02955.

XX 15-MAR-2000; 2000DE-1013847.

XX 06-APR-2000; 2000DE-1019058.

XX 07-APR-2000; 2000DE-1019173.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPISNOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2001-602752/68.

XX Claim 1; SEQ ID NO 236; 27bp; English.

XX The invention relates to a nucleic acid comprising a sequence of 18  
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with  
CC bisulphite, of genes associated with tumour suppression and  
CC oncogenes having a sequence taken from 536 (actually 533 since  
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences  
CC (58) and sequences complementary to (58). The nucleic acid may be a  
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may  
CC form part of a set of probes for detecting the cytosine methylation state  
CC and/or single nucleotide polymorphisms and also to be used in an  
CC array for analysing diseases associated with CpG dinucleotides e.g.  
CC cancers and tumours. The probes can also be used in a method for  
CC ascertaining genetic and/or epigenetic parameters for the diagnosis  
CC and/or therapy of existing diseases or the predisposition to specific  
CC diseases, by analysing cytosine methylations. The parameters may be  
CC compared to another set of genetic and/or epigenetic parameters, the

CC differences serving as basis for diagnosis and/or prognosis events which  
CC are disadvantageous to patients. The present sequence is one of the  
CC 533 genomic sequences derived from tumour suppressor genes and  
CC oncogenes. Sequences with even numbered Seq ID numbers are the  
CC complementary sequence of the corresponding odd numbered sequence (e.g.  
CC ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence  
CC is missing).  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 5297 BP; 1353 A; 107 C; 1126 G; 2711 T; 0 other;

Query Match	52;	8.4%;	Score	52;	DB	22;	Length	5297;	
Best Local Similarity	46.6%;		Pred.	No.	0.014;				
Matches	203;	Conservative	0;	Mismatches	230;	Indels	3;	Gaps	1

Qy	158	GTCAATTTGAAACAATTTCACATAATTTTACAGATCAATTCGGTTGAGGTACG	217
Db	2669	GTAAATTTTGATTTGGTTTGGTTTACGTTTATATATGATTAAGTATTAATTTAGST	2728
Qy	218	CACAGTCAGTAA--TGCTTATATTAATTAAGTGTCAAGAGATACCTAATTTAATGA	274
Db	2729	TATATTTGGAAATAGAGAGTTTACGTTTTCGTATGATGTTTTTTGAAATTTTGGAGTT	2788
Qy	275	TGCCAATTTACCTTTTATTTGCTGTACGGCATGATTTCAATTAAGTGTACAGTTGGTG	334
Db	2789	GTATGTTTTTATTTAATTTTGTATATAGGAATTTGTTTTAATTTTGTTCGAGATACGA	2848
Qy	335	CATTATGACATCAAGAGTTTAAATTTAAATGAGATAAATTAATTAATTAAGAGGT	394
Db	2849	AAATTTAAATTTGTTTTGTATATTTTAAATTTGTTAAATATTTTGCATTTAGAGAAATTT	2908
Qy	395	TGAAAGCAATTTGCTCTTCACAAACAACATTTGTCACTTTTACGTAGTTAGTTCAG	454
Db	2909	TTATATATGAATAATGTTTTTATAGTAGAATTAAGTTCGGTGTATTAATTAATTTAAT	2968
Qy	455	TAATTTGTAATAGGTAATGTTCACATATATGATTAATAAAGAGAGTTTCAACTTTTAC	514
Db	2969	TTGATATTTGTTAATTTAATTTATATAGTTAAATTTTAAATTAAGGTTTATTTTAA	3028
Qy	515	CATTATATTAATGCAAGTCCTTCAGGTGTGGCAGATTAATATGCTTAATACAGGAATAGAC	574
Db	3029	GATTTTACGTTGATGTATTTAATTTTATTTTGTATGATTAGTTAATAATTAAGGTTTTAT	3088
Qy	575	TTGTTTTATATACGT 590	
Db	3089	ATTTTTTTTTAGGTT 3104	

```

RESULT 11
ABL33559
ID  ABL33559 standard; DNA; 7128 BP.

```

AC ABL33559;

DT 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 1532

KW Human immune system disease; cytosine methylation; antiasmatic;  
KW antiarteriosclerotic; antinaemic; cytosolic; nootropic;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmologic;  
KW antirheumatic; antiarthritic; antidysrhythmic; antiparasitic;  
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
KW gene; ds.

OS	Homo sapiens.
XX	
PN	WO200200928-A2

XX 03-JAN-2002.  
PD  
XX  
PF 02-JUL-2001; 2001WO-EP07537.  
XX  
XX 30-JUN-2000; 2000DE-1034529.  
PR 01-SEP-2000; 2000DE-1043826.  
XX  
XX (EPIC-) EPIGENOMICS AG.  
XX  
XX  
PI Olek A, Piepenbrock C, Berlin K;  
XX  
DR WPI; 2002-130909/17.

PT Nucleic acid comprising fragment of chemically modified gene, useful  
PT for diagnosis and treatment of diseases associated with abnormal  
PT cytosine methylation -  
v v

PS Claim 1; SEQ ID NO 1532; 32bp + Sequence Listing; German

The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines. The sequences can be used in the diagnosis and treatment of immune system disorders, including eye diseases such as retinopathy, neovascular glaucoma and macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis, rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel diseases. The present sequence is a gene of the invention.

SQ Sequence 7128 BP; 1875 A; 207 C; 1610 G; 3436 T; 0 other;

Query Match	52;	8.4%;	Score	52;	DB	24;	length	7128;	
Best Local Similarity	46.5%;								
Matches	240;	Conservative	0;	Mismatches	270;	Indels	6;	Gaps	2

QY	78	TAATCAGAGAGGTTACTAAAGCATTCACACGACGAGGAGTGCGGGCTTGCAAT	137
Db	4809	TAAATTTATATAGATTATTTAAATGTATTTTATAGTTTAAATGTATAATGTATTTAAT	4868
QY	138	TTATTCAGGCGTAATGGAGCGTCATTTTGAACAATTTTCTACTATATTTTTCAGATC	197
Db	4869	TTAGTTAGTGTATGATGAATTTATTTTGTATTTATTTTATTAATAATTTTATTAATAAT	4928
QY	198	ATTTCGGTTTGAGGTTACAGACAGCAGATATAGCTTTATTTATTTATGTGCTCAAGA	257
Db	4929	TTTAAATATGATATTAATAATTAAGATTAATAATAGATTATATGATTAGTTTTCATGTTT	4988
QY	258	GATACGTAATTTATTTAGTATGCCAATAATTACTTTTATTCGCTGTACGGCAGATTTCAAT	317
Db	4989	TAT-----TAAGATGAGAAATATTTGAAAGCTTTTTTTTATGAGATTAATTTAGTAATTT	5043
QY	318	ACGTGTCAGATGTCGTCATATATGAGACATCAAAAGTTTTPAATTTAAATGAGATTAAT	377
Db	5044	TAGTTTATATTTTAGTTATTTATTTATTTATTTAATAA-TTATTTATGAGATTTAATTTATG	5102
QY	378	TAATATATATAAAGGTTGAAGAAGATGTTGCTTCCACAAACACTGTTGCACTTTT	437
Db	5103	TTAAGATATGTTTATAGTTATTTAGGAGATATAAAGATATATATATATATATGTTAATTTAT	5162
QY	438	ACGTAGTTTACGTCAAGTAATTTTATAGTATTTGTTCCATATATGATTAATAAAGAGA	497
Db	5163	GGGGTTTATATTTTAGTGTAATAATTTATGATTTTAAAGTTAAATTAATGAGAGAAATAT	5222
QY	498	GTTTTCAAACTTTTACCATATATATATGCAAGTCCTCAGGTGCGGAGATTAATGCT	557
Db	5223	GATTTTATGCTTTTGTTTTATTAAGATATGTTATGGAAGTATAGTATATGATTT	5282
QY	558	TAATACAGATATGACTGTTTATATATACGCTAAT	593
Db	5283	TAAATTTTAAATATTAATTAATGATTAATTTATGTAAT	5318

## RESULT 12

AB067150	standard; DNA; 40324 BP.
AB067150;	
28-AUG-2002	(first entry)
Human angiogenesis associated polynucleotide SEQ ID NO 180.	
Human; angiogenesis; methylation; eye disease; glaucoma; tumour;	
inflammation; rheumatoid arthritis; diabetic retinopathy; antiulcers;	
macular degeneration; inflammatory bowel disease; Crohn's disease;	
antirheumatic; antidiabetic; antidiabetic; antidiabetic;	
arteriosclerotic; ds.	
Homo sapiens.	
MO200246454-A2.	
13-JUN-2002.	
06-DEC-2001; 2001MO-EP14320.	
06-DEC-2000; 2000DE-1061338.	
(EPIC-) EPIGENOMICS AG.	
Schacht O;	
WPI; 2002-500450/53.	
New nucleic acid fragments from chemically treated	
angiogenesis-associated genes, useful for determining methylation	
status, e.g. in diagnosis or treatment of cancer	
Claim 1; SEQ ID NO 180; 41bp + Sequence Listing; German.	
The invention relates to a nucleic acid (I) comprising a segment of 18	
bases of chemically pretreated DNA of angiogenesis-associated genes (II)	
having sequences (AB066971-AB067178) or their complements. (I), also	
related oligomers, are used to evaluate the methylation status and/or	
single-nucleotide polymorphisms, in angiogenesis-related genes, for	
diagnosis and treatment of eye diseases, proliferative retinopathy,	
neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,	
diabetic retinopathy, macular degeneration caused by neovascularisation,	
psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and	
Crohn's disease.	
Note: The sequence data for this patent did not form part of the printed	
specification, but was obtained in electronic format directly from WIPO	
at <a href="http://wipo.int/pub/published_pct_sequences">ftp://wipo.int/pub/published_pct_sequences</a> .	
Sequence 40324 BP; 12266 A; 436 C; 7652 G; 19970 T; 0 other;	
Query Match	8.4%; Score 52; DB 24; Length 40324;
Best Local Similarity	46.5%; Pred. No. 0.022;
Matches 240; Conservative 0; Mismatches 270; Indels 6; Gaps 2;	
78	TTAATCAGAAGAGGTTACTAAAGATTGACTACTGACGACGAGGATGCTGGGCTTCTAT 137
38005	TAAATTTATAGATTATTTAATTGTAATTTAAAGTTTAAATGTAATGATGTAATAT 38066
138	TTATTCAGGCGTATAGGAGCGTCAATTTTGAACAATTTTCTATATATTTTACAGAAAC 197
38065	TTACTGTAGTATGATGTAATTTATTTTGTATTTATTTATTTATTTAATTTTAAATAAT 38122
198	ATTTCGGTTGAGGTTACAGACACAGTCAGTAATGCTTATTTATTTAATGTTGCTCAAG 257
38125	TTTAAATATATGATTTAAATTAAGTTAATATGATTAATGATTAATGTTTATGTTATTT 38184
258	GATAGCTATTTTATTTATGATGCAATATATACCTTTTATTTGCGTTACGGCAGTTTCAT 317
38185	TAT-----TATGATGAGAAATATTTTAAAGTTTATTTTATGAGATTAATTAATTT 38233

Oy	318	ACGGTCAAGGTGGGCAATTATGACTCAAAAGCTTTTAATTAAATGAGATAATT	377
Dd	38240	TAGTTTATATTTTAGTTATTTATTTATTAATA-TTATTTATAGTATTTAATTATG	38298
Oy	378	TAATATATATAAAGGGTTGAAGAATGTTTGCTCTCAACAACACTGTCGACTTT	437
Dd	38299	TTAGATATTTGTTTTAGTTATTTAGGATATAAAGATATATATATATGTTATATTTAT	38356
Oy	438	ACGTAGTTTAGTCCAAATATTTGTTATAGTATTTGCCATATATGATTTATAAAGACA	497
Dd	38359	GGGGTTTATATTTTATGTTGTAATATTTGATTTTAAGTGTAAATATATGAGAAATAT	38418
Oy	498	GTTTTCAAACCTTTTTCACATTAATATATGCAAGTCCCTCAGGTCGCGCATATATGCT	557
Dd	38419	GATTTTATGCTTTTGTGTTTTTTTGAATGATTTGTATAGGAATATGTTATATGATTT	38478
Oy	558	TAATACAGAAATAGTACTGTTTATATATACGCTAAT	593
Dd	38479	TAATATTTAATATATTAATTAATGATATTTATGATTT	38514

  

XX	ABL32164	standard; DNA; 5523 BP.
XX	ABL32164;	
XX	26-MAR-2002	(first entry)
DT		
XX		
DE		Human immune system associated gene SEQ ID NO: 137.
KW		Human; immune system disease; cytosine methylation; antiaesthetic;
KW		antiathectosclerotic; antiaiemic; cytosatic; nootropic;
KW		neuroprospective; anti-HIV; anticonvulsant; ophthalmological;
KW		antiinematoc; antiarthritis; antidiabetic; antipsoriatic;
KW		antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW		acute myeloid leukemia; Alzheimer's disease; AIDS; epilepsy;
KW		neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
XX		gene; ds.
OS		Homo sapiens.
XX		
PN	WO200200928-A2.	
XX		
PD	03-JAN-2002.	
XX		
PF	02-JUL-2001; 2001MO-EP07537.	
XX		
PR	30-JUN-2000; 2000DE-1032529.	
XX		
PR	01-SEP-2000; 2000DE-1043826.	
XX		
PA	(EPIG-) EPIGENOMICS AG.	
XX		
PI	Olek A, Piepenbrock C, Berlin K;	
DR	WPI, 2002-130909/17.	
PT		Nucleic acid comprising fragment of chemically modified gene, useful
PT		for diagnosis and treatment of diseases associated with abnormal
PT		cytosine methylation -
XX		
PS	Claim 1; SEQ ID NO 137; 32pp + Sequence Listing; German.	
XX		
CC		The present invention provides a number of human immune system associated
CC		genes which are modified by the methylation of cytosines. The sequences
CC		can be used in the diagnosis and treatment of immune system disorders,
CC		including eye diseases such as retinopathy, neovascular glaucoma and
CC		macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC		leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC		rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC		diseases. The present sequence is a gene of the invention.
XQ	Sequence 5523 BP, 1378 A, 119 C, 1366 G, 2660 T, 0 other;	



XX 03-JAN-2002.  
 PD  
 XX  
 PF 02-JUL-2001; 2001WO-EP07537.  
 XX  
 XX 30-JUN-2000; 2000DE-1032529.  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX  
 PA (EPIC-) EPIGENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K;  
 DR WPI; 2002-130909/17.  
 XX  
 XX Nucleic acid comprising fragment of chemically modified gene, useful  
 PT for diagnosis and treatment of diseases associated with abnormal  
 PT cytosine methylation -  
 XX  
 PS Claim 1; SEQ ID NO 2097; 32pp + Sequence Listing; German.  
 XX  
 CC The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
 CC diseases. The present sequence is a gene of the invention.  
 CC  
 XX  
 SQ Sequence 73334 BP; 18968 A; 858 C; 15329 G; 38179 T; 0 other;  
 Query Match 8.3%; Score 51.8; DB 24; Length 73334;  
 Best Local Similarity 49.8%; Pred. No. 0.028;  
 Matches 133; Conservative 0; Mismatches 132; Indels 0; Gaps 0;  
 QY 264 TATTATATGATGCCAATATCTTTTATGCTGTACGCGATGATTCATTACGTGT 323  
 DB 57666 TAAGTATTTAATGGAATTAAGAATTTTATGATGATTTTAAAAAAGTTA 57725  
 QY 324 ACAAGTGTGCTATGACCTACAAAGCTTTTAAATTTAAATGAGTAAATTAAT 383  
 DB 57726 TAAATATAGGGGTAATATGATTTATTTTGTGAGGAACTAAAGAAATTAAT 57785  
 QY 384 AATTAAGGCTGTAAGAAAGATGTTGCTCTCAACAACACTTGTGACCTTTACGTAG 443  
 DB 57786 TTTAGTAAATTAATAATAGTTATAGTTTATTAAGAAAGAAATTTTGTAAAT 57845  
 QY 444 TTTAGTCAAGTATGTTATAGTATGTTCCATATATGATTAATAAGAGAGTTTC 503  
 DB 57846 TTAAGTGAATGAAATTAATTTTAAATTTATTTGATTAATTTTAAAGGGGTTT 57905  
 QY 504 AAACCTTTTACATTAATATATG 526  
 DB 57906 AAATATTTTGTGATTAAGTTATG 57928  
 RESULT 16  
 ABL33299  
 ID ABL33299 standard; DNA; 12025 BP.  
 AC ABL33299;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Human immune system associated gene SEQ ID NO: 1272.  
 XX  
 XX Human; immune system disease; cytosine methylation; antiasthmatic;  
 KW antiasthmatic; anti-HIV; anticonvulsant; ophthalmological;  
 KW antineumatic; antiarthritic; antidiabetic; antipsoriatic;  
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;

KW gene; ds.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200200928-A2.  
 PN  
 XX  
 XX 03-JAN-2002.  
 PD  
 XX  
 XX 02-JUL-2001; 2001WO-EP07537.  
 PF  
 XX 30-JUN-2000; 2000DE-1032529.  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX  
 PA (EPIC-) EPIGENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K;  
 DR WPI; 2002-130909/17.  
 XX  
 XX Nucleic acid comprising fragment of chemically modified gene, useful  
 PT for diagnosis and treatment of diseases associated with abnormal  
 PT cytosine methylation -  
 XX  
 PS Claim 1; SEQ ID NO 1272; 32pp + Sequence Listing; German.  
 XX  
 CC The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
 CC diseases. The present sequence is a gene of the invention.  
 CC  
 XX  
 SQ Sequence 12025 BP; 4009 A; 84 C; 2126 G; 5806 T; 0 other;  
 Query Match 8.2%; Score 51.2; DB 24; Length 12025;  
 Best Local Similarity 46.5%; Pred. No. 0.026;  
 Matches 201; Conservative 0; Mismatches 228; Indels 3; Gaps 1;  
 QY 163 TTGAAACAATTTTCTACTATATTTTACAGAAATCTGGTTGAGGTTACAGACAG 222  
 DB 203 TTTTACAGAAATTTTACGTTATTTTAAATTTTATGATGATGTTGATTA 262  
 QY 223 TCAGTATATGCTTATTTATATGTTGCTCAAGATAGCTATTTATGATGCCAAT 282  
 DB 263 TAGATATTTGGGTTTATTTTGTAAAGTTATTTTATTTTAAATATATGAAAT 322  
 QY 283 TTAATTTTATGCTGTACGCGATGATTTCAATGATCAAGTGTGATTAATG 342  
 DB 323 TGAATTTAAGTAGAGATTTATTTAGAAATTTGTTGATTAATATGATTAATTCGTA 382  
 QY 343 ACTCAAGAGTTTAAATTTAAATGAGTAATTTAATATTAATAAGGTTGAAGGA 402  
 DB 383 ATTAAGCTATATTATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 442  
 QY 403 AATGTTGCTTCAACAACACTTG--TTCACTTTACGATGTTAGTCAAGTAAT 459  
 DB 443 AAGTTTATTAATTTGATATTTTGTATTAAGAAATTTTATTAAGTAAATTTATTT 502  
 QY 460 GTTATAGTATTTGCTCATATATGATTAATAAGAGAGTTTCAACTTTTACCATTA 519  
 DB 503 TATTATGTAATTTATTAATTTTATGATTAAGTAATTTTATGCTGATTAATGTA 562  
 QY 520 TATTATGCAAGTCTTCAGGTGCGAGATTATAGCTTAATACAGAAATGACTTGT 579  
 DB 563 TTTTAATTAAGTGTGCTGGAATTAATAATATGTTTATTAATGAATTTTACTATAGGT 622  
 QY 580 TTTATATACGCTA 591  
 DB 623 TTTAAGTTGGTA 634

ID	ABL33031	standard; DNA; 6118 BP.
AC	ABL33031;	
AD	26-MAR-2002	(first entry)
AE		
AF	Human immune system associated gene SEQ ID NO: 1004.	
AG		
AH	Human, immune system disease; cytosine methylation; antiasthmatic;	
AI	antiartherosclerotic; antianaemic; cytosolic; nocrotic;	
AJ	neuroprotective; anti-HIV; anticonvulsant; ophthalmological;	
AK	anthelmatic; antiarthritic; antidiabetic; antipsoriatic;	
AL	antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;	
AM	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;	
AN	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;	
AO	gene; ds.	
AP	Homo sapiens.	
AQ	WO200200928-A2.	
AR	03-JAN-2002.	
AS	02-JUL-2001; 2001WO-EP07537.	
AT	30-JUN-2000; 2000DE-1032529.	
AV	01-SEP-2000; 2000DE-1043826.	
AW	(EPIC-) EPICENOMICS AG.	
AX	Olek A, Piepenbrock C, Berlin K;	
AY	WPI; 2002-130909/17.	
AZ	Nucleic acid comprising fragment of chemically modified gene, useful	
BA	for diagnosis and treatment of diseases associated with abnormal	
BB	cytosine methylation -	
BC	Claim 1; SEQ ID NO 1004; 32bp + Sequence Listing; German.	
BD	The present invention provides a number of human immune system associated	
BE	genes which are modified by the methylation of cytosines. The sequences	
BF	can be used in the diagnosis and treatment of immune system disorders,	
BG	including eye diseases such as retinopathy, neovascular glaucoma and	
BH	macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid	
BI	leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,	
BJ	rheumatoid arthritis, psoriasis and inflammatory/alcerative bowel	
BK	diseases. The present sequence is a gene of the invention.	
BL	Sequence 6118 BP; 1751 A; 141 C; 1229 G; 2997 T; 0 other;	
BM	Query Match 8.2%; Score 51; DB 24; Length 6118;	
BN	Best Local Similarity 50.5%; Pred. No. 0.025;	
BO	Matches 151; Conservative 0; Mismatches 145; Indels 3; Gaps 1;	
BP	227 TATATGCTTATTTATTTATGTTGCTCAAGATAGCATTTTATGATGCAATATAC 286	
BQ	4654 TAATTACGTTATTTTATTTTATTTATTAATTAATTAATGTTAAAGTTATGTTATTTT 4713	
BR	287 TTTTATTTGCTGTATACGCGATGATTTGATTACGTTACGTTCAAGTTGGTCATTATGACTA 346	
BS	4714 TTTTATTTGAGGTTTGTGTTTAANT---TTATTTAAGAAGTTTATGTTTATTTT 4770	
BT	347 CAAAGCTTTTAAATTTAAATGAGTAATTTAATATAATAAAGGTTGAAAGAAATGT 406	
BU	4771 GAGATTTTAAATAATGAATAATGAATAATATATAGCTTATGATGATGAGTTAAGCAATTG 4830	
BV	407 TTGCTTCGAACAACAACCTTTGTGACTTTTACGATTTTGGTTCATTAATTTGTTATAG 466	
BW	4831 AAGTTTCATTTATTTGGCATTAAGTTTATATATATGAGTTAAGTTTATTTATTTT 4890	



XX WPI; 2002-130909/17.  
 DR Nucleic acid comprising fragment of chemically modified gene, useful  
 XX for diagnosis and treatment of diseases associated with abnormal  
 PT cytosine methylation  
 XX  
 PS Claim 1; SEQ ID NO 1804; 32pp + Sequence Listing; German.  
 XX  
 CC The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
 CC rheumatoid arthritis, psoriasis and inflammatory/nutritional bowel  
 CC diseases. The present sequence is a gene of the invention.  
 XX  
 SQ Sequence 6126 BP; 1772 A; 88 C; 1318 G; 2948 T; 0 other;  
 Query Match 8.1%; Score 50.6; DB 24; Length 6126;  
 Best Local Similarity 46.7%; Pred. No. 0.031;  
 Matches 161; Conservative 0; Mismatches 184; Indels 0; Gaps 0;  
 QY 203 GGTGGAGTTACACACAGTCAGTATGCTTTATTTATTTATTTGCTCAGAGATAG 262  
 DB 70 GGTGAGTTTAAATATTTATTTTAAATTTTAAATATTTATTTGCTTTAAATATGCT 129  
 QY 263 CTATTTTATGATGCCAATTTACTTTTATGCTGTTACGCGATGATTCATTACGTG 322  
 DB 130 TTTTATTTGTTTGTATTAAGTTTGAATTTTATTTATTTATTTATTTATTTAGTGT 189  
 QY 323 TACAAAGTTGTCATATGATGACACAAAGTTTAAATTTAAATGAGTAAATTTATA 382  
 DB 190 ATAGAGAAATTAAGTTATTTGCTTATGAAAAATTTATTTGTAATTTTATTTATTTGAT 249.  
 QY 383 TAATTAAGAGGTGAAGAAATGTTGCTTCAACAAACTGTTGACCTTTACGTA 442  
 DB 250 AAAATTTAAGTTTAAAAAAGTTAAGTTATAGATTTTATTTTATTTTAAATATAT 309  
 QY 443 GTTTAGTTCAAGTAATTTGTATAGTATTTGTCATATATGATTTAAAGAGAGTTT 502  
 DB 310 TTTAAATTAATTTTATTTAGTTTATTTAGTGAATTTTAAATTTTATTTAGATTTAGAT 369  
 QY 503 CAACCTTTTACATTTATTTATGCAAGTCCCTCAGGCTGGCG 547  
 DB 370 ATAAATATTTGTTAATATTTTATTTATGATTAATGAGGTTGAGAG 414  
 RESULT 21  
 AAS45480  
 ID AAS45480 standard; DNA; 11047 BP.  
 XX  
 AC AAS45480;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Chemically pretreated genomic DNA associated with cell cycle #93.  
 XX  
 KW Cell cycle; human; CpG dinucleotide; cytosine methylation; HIV; aging;  
 KW human immunodeficiency virus; neurodegenerative disorder; solid tumour;  
 KW graft-versus-host disease; glomerular disease; Lewy body disease; cancer;  
 KW arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritis;  
 KW immunosuppressive; antitumour; cytosolic; antiarteriosclerotic; ds;  
 KW PCR primer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200168911-A2.  
 XX  
 PD 20-SEP-2001.  
 XX  
 PF 15-MAR-2001; 2001WO-EP02945.

XX 15-MAR-2000; 2000DE-1013847.  
 PR 06-APR-2000; 2000DE-1019058.  
 PR 07-APR-2000; 2000DE-1019173.  
 PR 30-JUN-2000; 2000DE-1032529.  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX  
 PA (EPIC-) EPIGENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K;  
 XX  
 DR WPI; 2001-602751/68.  
 XX  
 PT Designing primers and probes for analysing diseases associated with  
 PT cytosine methylation state e.g. arthritis, cancer, aging,  
 PT arteriosclerosis comprising fragments of chemically modified genes.  
 PT associated with cell cycle -  
 XX  
 PS Claim 1; SEQ ID No 185; 28pp; English.  
 XX  
 CC Sequences AAS45296-AAS45520 represent chemically pretreated genomic DNA  
 CC molecules associated with the cell cycle and specific PCR primers of the  
 CC invention. The sequences are useful for detecting the methylation state  
 CC of all CpG dinucleotides in a sequence and therefore for analysing  
 CC associated diseases. By analysing cytosine methylations in the pretreated  
 CC DNA, genetic and/or epigenetic parameters for the diagnosis and therapy  
 CC of existing diseases or the predisposition to specific diseases can be  
 CC ascertained. The parameters may be compared to another set of genetic  
 CC and/or epigenetic parameters, the differences serving as basis for  
 CC diagnosis and/or prognostic events which are disadvantageous to patients.  
 CC The sequences of the invention are useful for the diagnosis and therapy  
 CC of HIV infection, neurodegenerative disorders, graft-versus-host disease,  
 CC aging, glomerular disease, Lewy body disease, arthritis,  
 CC arteriosclerosis, solid tumours and cancers.  
 XX  
 SQ Sequence 11047 BP; 3237 A; 248 C; 2358 G; 5204 T; 0 other;  
 Query Match 8.1%; Score 50.6; DB 22; Length 11047;  
 Best Local Similarity 46.9%; Pred. No. 0.036;  
 Matches 158; Conservative 0; Mismatches 179; Indels 0; Gaps 0;  
 QY 164 TTGAACAATTTCTACTATTTATTTTACAGAAATCATTTGGTTGAGTTACGACAGT 223  
 DB 10382 TTGAATGAATTTAGATTAATTTATTTGAATTAATTTTGTAGTTGTTAGTAGA 10441  
 QY 224 CAGTATATGCTTTATTTATTTATTTGCTCAAGATAGCTATTTATTTATTCACATAT 283  
 DB 10442 GCGATTAATGTTATTTAGTATTAATTTGTTGAATTTTATTTTATTTGTTGAAGTTGG 10501  
 QY 284 TACTTTTATTTGCTTTACGGCATGATTTTCATTAGTGACAAAGTTGGCATTTATGGA 343  
 DB 10502 TTTTATTTTAAAGTAATGTAATTAAGAAATTTAGATTAATTTGAGTTTATTTT 10561  
 QY 344 CTACAAAGTTTAAATTTAAATGAGTAATTTAATTAATTAATTAAGGTTGAAAGAA 403  
 DB 10562 TTGATTAATTTTATTTTAAATTTATTTATTTTTCGTTAAGAAAGTTGTAATTAATTA 10621  
 QY 404 TGTTCCTTCTCAACAACACTGTTGCACTTTTACGTAGTTTATTTCAAGTAATTTGTTA 463  
 DB 10622 TTTTATTTTAAATTTGTAATTTGTTTGTGTTATGATGAGTAATTTAGTTATTTGTTT 10681  
 QY 464 TAGGATTTGTCATATGATTAATTAAGAGAGAGTT 500  
 DB 10682 TTGATTAATTTTATTTGATTTTATTTAGTTATTTAGTTGGAATTT 10718  
 RESULT 22  
 ABL33985  
 ID ABL33985 standard; DNA; 11047 BP.  
 XX  
 AC ABL33985;  
 XX  
 DT 26-MAR-2002 (first entry)



XX DE Human immune system associated gene SEQ ID NO: 1958.  
 XX XX  
 KW Human; immune system disease; cytosine methylation; antiasthmatic;  
 KW antiarteriosclerotic; antianemic; cytoseric; noctropic;  
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
 KW antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;  
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
 KW gene; ds.  
 XX XX  
 OS Homo sapiens.  
 XX XX  
 PN WO200200928-A2.  
 XX XX  
 PD 03-JAN-2002.  
 XX XX  
 PF 02-JUL-2001; 2001WO-EP07537.  
 XX XX  
 PR 30-JUN-2000; 2000DE-1032529.  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX XX  
 PA (EPig-) EPIGENOMICS AG.  
 XX XX  
 PI Olek A, Piepenbrock C, Berlin K;  
 DR WPI; 2002-130909/17.  
 XX XX  
 PT Nucleic acid comprising fragment of chemically modified gene, useful  
 PT for diagnosis and treatment of diseases associated with abnormal  
 PT cytosine methylation -  
 XX XX  
 PS Claim 1; SEQ ID NO 1958; 32pp + Sequence Listing; German.  
 XX XX  
 CC The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
 CC diseases. The present sequence is a gene of the invention.  
 XX XX  
 SQ Sequence 11047 BP; 3237 A; 248 C; 2358 G; 5204 T; 0 other;

Query Match 8.1%; Score 50.6; DB 24; Length 11047;  
 Best Local Similarity 46.9%; Pred. No. 0.036;  
 Matches 158; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY 164 TTGAACAAATTTCTACTATATTTTACAGATCATTTGCGTTGAGGTTACGACACACT 223  
 DB 10382 TTGAACAAATTTAGATTAATTTTGAAGATTAATTAATTTTGAATTTGTTAGTAGAGA 10441  
 QY 224 CAGTAATGCTTAATTTATTTATTTGCTCAAGATGAGTATTTTATGATGCAATAT 283  
 DB 10442 GCGATTAATTTGTTATTTGTTATTTGTTGATTTTATTTTGTGTTGAAGTTGG 10501  
 QY 284 TACTTTTATTTGCTTTAGCGCATGATTTTCATTAGCTGATCAAGTTGGTCATTATGA 343  
 DB 10502 TTTTATTTTAAAGTAATGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTTGTTT 10561  
 QY 344 CTACAAAGCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 403  
 DB 10562 TTGATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTAA 10621  
 QY 404 TGTTCCTTCACAAACACTTGTGACCTTTTACAGTATTTAGTTCAAGTAATTTGTA 463  
 DB 10622 TTTTATTTTATTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 10681  
 QY 464 TAGGTATTTGTTCAATATATGATTAATTAAGAGAGAGTT 500  
 DB 10682 TTGATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 10718

RESULT 23  
 ID ABRK28402 standard; DNA; 11047 BP.  
 XX ABRK28402;  
 AC ABRK28402;  
 XX XX  
 DT 23-APR-2002 (first entry)  
 XX XX  
 DE DNA transcription associated complementary genomic DNA #138.  
 XX XX  
 KW DNA transcription associated gene; peptide nucleic acid; PNA-oligomer;  
 KW PNA; cytosine methylation state; SNP; retroviral infection; gene; ds;  
 KW single nucleotide polymorphism; adenosine deaminase deficiency; cancer;  
 KW viral infection; Sezary syndrome; haematological disorder; tuberculosis;  
 KW immunological disorder; Werner syndrome; developmental disorder;  
 KW psoriasis; Rieger's syndrome; neurological disorder; erythropoiesis;  
 KW neurodegenerative disorder; Wardenburg syndrome; Niemann-Pick disease;  
 KW myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;  
 KW angiogenesis; congenital heart disease; HDR syndrome; gene therapy;  
 KW polyglutamine disorder; solid tumour.  
 XX XX  
 OS Unidentified.  
 XX XX  
 PN WO200192565-A2.  
 XX XX  
 PD 06-DEC-2001.  
 XX XX  
 PF 06-APR-2001; 2001WO-EP03973.  
 XX XX  
 PR 06-APR-2000; 2000DE-1019058.  
 PR 07-APR-2000; 2000DE-1019173.  
 PR 30-JUN-2000; 2000DE-1032529.  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX XX  
 PA (EPig-) EPIGENOMICS AG.  
 XX XX  
 PI Olek A, Piepenbrock C, Berlin K;  
 DR WPI; 2002-090046/12.  
 XX XX  
 PT New nucleic acids or oligomers, useful for diagnosing or treating  
 PT diseases associated with DNA transcription, e.g. immunological  
 PT disorders, Werner syndrome, psoriasis, myocardial infarction, solid  
 PT tumours or cancer -  
 XX XX  
 PS Claim 1; SEQ ID No 276; 32pp; English.  
 XX XX  
 CC The invention relates to a nucleic acid, which comprises a segment of the  
 CC chemically pretreated DNA of genes associated with DNA transcription from  
 CC one of 346 sequences, and an oligomer, in particular an oligonucleotide  
 CC or peptide nucleic acid (PNA)-oligomer that hybridises to or is identical  
 CC to the chemically pretreated DNA of genes associated with DNA  
 CC transcription. The set of oligomer probes are useful for detecting the  
 CC cytosine methylation state and/or single nucleotide polymorphisms (SNPs)  
 CC in a chemically pretreated genomic DNA. The nucleic acids are useful for  
 CC diagnosing or treating diseases associated with DNA transcription  
 CC (particularly with the methylation status), e.g. adenosine deaminase  
 CC deficiency, viral infection, retroviral infection, Sezary syndrome,  
 CC haematological disorders, immunological disorders, Werner syndrome,  
 CC tuberculosis, developmental disorders, psoriasis, Rieger's syndrome,  
 CC neurological disorders, neurodegenerative disorders, Wardenburg  
 CC syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial  
 CC infarction, hypertension, angiogenesis, erythropoiesis, congenital heart  
 CC disease, HDR syndrome, arthritis, polyglutamine disorders, solid tumours  
 CC or cancer. Sequences ABRK28127-ABK28472 represent DNA transcription  
 CC associated genomic DNA molecules of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format directly from the  
 CC European Patent Office.  
 XX XX  
 SQ Sequence 11047 BP; 3237 A; 248 C; 2358 G; 5204 T; 0 other;



PT Novel nucleic acid useful for diagnosis and therapy of diseases  
PT associated with development genes such as diabetes, comprises a  
PT sequence of a segment of chemically pretreated DNA of genes associated  
PT with development -

PS Claim 1; SEQ ID NO 65; 27pp; English.

XX  
XX The invention relates to a nucleic acid (I) comprising a sequence at  
CC least 18 bases in length of a segment of chemically pretreated DNA (II)  
CC of genes associated with development selected from 87 genes listed in  
CC the specification such as ACCPN, ADPN, or APD1 and comprising one of 350  
CC sequences (ABN7984-ABN8033) or their complements. The invention is  
CC useful for the diagnosis or therapy of diseases associated with  
CC development genes, in particular diseases related to homeobox containing  
CC genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes  
CC associated with congenital heart disease, epilepsy, diseases related to  
CC histone deacetylation, Currarino syndrome, diseases related with the  
CC development of the brain and limb girdle muscular dystrophy and dwarfism.  
CC Oligomers specific to each of the genes are useful for detecting the  
CC methylation state of all CpG dinucleotides within the 350 sequences or  
CC (II) and their complementary sequences, (II) and/or their complements and  
CC the amplification of the 350 sequences, (II) and/or their complements and  
CC as oligomer probes for detecting the cytosine methylation state and/or  
CC single nucleotide polymorphisms (SNPs).  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification but is based on sequence information supplied to Derwent by  
CC the European Patent Office.

XX  
XX Sequence 9543 BP; 2652 A; 145 C; 1975 G; 4771 T; 0 other;

Query Match 8.1%; Score 50.2; DB 24; Length 9543;  
Best Local Similarity 46.2%; Pred. No. 0.043;  
Matches 166; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

QY 236 TATTATTATTTGTCGACAGATFACCTATTATTGAGCCAAATTAATTATTG 235  
DB 2627 TAAGTTTATTTTATTTTTCGTAATTTTAAAGTTTGGTAATTAATTTTTCG 2686  
QY 296 CTGTACGATGATTTCAATTAAGTACAGTGTGATTAATGATTAAGTAAAGTT 355  
DB 2687 TGTTTATGATTTTATTTTATTTTATTTTATTTTAAAGTAAATTAATTTTAT 2746  
QY 356 TTTAAATTTAATGAGTAAATTTTAAATTTAATTTAATTTAATTTTATTTTAT 415  
DB 2747 TTTTATTTATGTTTGTATTTTATTTTATTTTATTTTAAAGTAAATTTTAT 2806  
QY 416 AACAAACACTGTTGCACTTTTACGATTTAGTTCAAGTAATTTGTTATGTTTC 475  
DB 2807 AGTATATATTTATTTTATTTTAAAGTTAATTAATTTTGTATTAATTTTATA 2866  
QY 476 CATATATGATTTAATTAAGAGAGTTTCAACCTTTTACCAATTAATTAATTAAGAG 535  
DB 2867 TATATATATATATATATATATATATATTTTGTATTTTATTTTATTTATGAT 2926  
QY 536 CAGGTGTGAGATTTATATGCTTAATACAGAAATGACTTTGTTTATATACGTAAT 594  
DB 2927 TGATTTGTTTATTTTATTTTGTATTTTATTTTATTTTGTATTTTGAATTA 2985

RESULT 26  
AA161373/c  
ID AA161373 standard; DNA; 513445 BP.

XX AA161373;

XX 16-OCT-2001 (first entry)

XX Soybean 318013 region A3, SEQ ID NO: 4.

XX Soybean; antihelminthic; gene therapy; soybean cyst nematode; SCN;  
KM SCN resistance; rhg1; Rhg4; SCN resistant allele; plant breeding;  
KM 240017 region G3; 318013 region A3; 515002 region G2; de.  
XX

OS Glycine max.

XX  
XX W0200151627-A2.

XX  
XX 19-JUL-2001.

XX  
XX 05-JAN-2001; 2001WO-US00552.

XX  
XX 07-JAN-2000; 2000US-0174880.

XX  
XX (MONS ) MONSANTO CO.

XX  
XX Hauge BM, Wang ML, Parsons JD, Parnell JD;

XX  
XX WPL; 2001-425872/45.

XX  
XX P-P-SDB; AAMA42216.

XX  
XX New purified nucleic acid for producing a soybean plant having soybean  
PT cyst nematode resistance and for use in plant breeding programs -

XX  
XX Claim 30; Page 596-893; 1353pp; English.

CC The invention relates to nucleic acid molecules from regions of the  
CC soybean genome which are associated with soybean cyst nematode (SCN)  
CC resistance. The nucleic acids are used to transform plants, and can  
CC produce soybean plants having an rhg1 or an Rhg4 SCN resistant allele.  
CC The nucleic acids can be used for investigating rhg1 or Rhg4 haplotypes  
CC of soybean plants and for introgressing SCN resistance or partial SCN  
CC resistance into soybean plants. They can also be used in plant breeding  
CC programmes. The invention also relates to proteins encoded by such  
CC nucleic acid molecules, as well as antibodies capable of recognising  
CC these proteins. The present sequence is a nucleic acid molecule  
CC provided in the specification.

XX  
XX Sequence 513445 BP; 173367 A; 85402 C; 83912 G; 170492 T; 272 other;

Query Match 8.1%; Score 50.2; DB 22; Length 513445;  
Best Local Similarity 47.4%; Pred. No. 0.099;  
Matches 213; Conservative 0; Mismatches 233; Indels 3; Gaps 2;

QY 147 CGTATGGAACGTCATTTGAAACAAATTTTCTACTAATTTTAA--CAGATCATTTCG 204  
DB 182573 CGTAGGGGGAATATGATTTTATCAAAATTTGTATGACTTATTTAAAGTATGATA 182514  
QY 205 TTGAGAGTTACAGACAGTCAGATATGCTTATTTATTTATTTGCTCAAGATAGCT 264  
DB 182513 ATTTTACTGGAATTAAGATATTTTAAACCTATTTTACTAATTTTCAAAATTCAAAAC 182454  
QY 265 ATTTTATTTATGCAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 324  
DB 182453 TATTATATATATATATATATATATATATTTTAAACCTAATTTTGAACCGGTTCTTACGAATA 182394  
QY 325 CAAAGTTGGATTAATGACCTCAAGAGTTTAAATTTAATTTAATTTAATTTAATTTAATA 384  
DB 182393 AAAAAGACATTTTATTTTACAGATTAATTTATTTTAAAGATTTCAAAAAGAAAAA 182334  
QY 385 ATAAAAGGTTGAAGAGATTTGCTTCTCAACAAACCTGTTGACCTTTTACGATG 444  
DB 182333 ATTAAGAAAATTTAAAGATTTTAAATTAAGAGAAAGATTTTGAATTAATTTTAT 182274  
QY 445 TTAGTTCAAGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 504  
DB 182273 TTTTATATTTGATTTATTTTCAAAAATTTAACTTA-ATTTTATTTTATTTTATTTTATTT 182215  
QY 505 AACTTTTACATTTATTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 564  
DB 182214 AATTATGCGTCAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 182155  
QY 565 GGAATAGTACTGTTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 593  
DB 182154 ATTTTATGCAAAATTTTCTCTATTTTATTTTATTTTATTTTATTTTATTTTATTT 182126

RESULT 27  
 AAS45426  
 ID AAS45426 standard; DNA; 5935 BP.  
 XX  
 AC AAS45426;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Chemically pretreated genomic DNA associated with cell cycle #66.  
 XX  
 KW Cell cycle; human; Cpg dinucleotide; cytosine methylation; HIV; aging;  
 KW human immunodeficiency virus; neurodegenerative disorder; solid tumour;  
 KW graft-versus-host disease; glomerular disease; Lewy body disease; cancer;  
 KW arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritis;  
 KW immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; ds;  
 KM PCR primer.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200168911-A2.  
 XX  
 PD 20-SEP-2001.  
 XX  
 PF 15-MAR-2001; 2001WO-EP02945.  
 XX  
 PR 15-MAR-2000; 2000DE-1013847.  
 PR 06-APR-2000; 2000DE-1019058.  
 PR 07-APR-2000; 2000DE-1019173.  
 PR 30-JUN-2000; 2000DE-1032529.  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX  
 PA (EPIG-) EPIGENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K;  
 XX  
 DR WPI; 2001-602751/68.  
 XX  
 PT Designing primers and probes for analysing diseases associated with  
 PT cytosine methylation state e.g. arthritis, cancer, aging,  
 PT arteriosclerosis comprising fragments of chemically modified genes.  
 PT associated with cell cycle -  
 XX  
 PS Claim 1; SEQ ID NO 131; 28pp; English.  
 XX  
 CC Sequences AAS45296-AAS45520 represent chemically pretreated genomic DNA  
 CC molecules associated with the cell cycle and specific PCR primers of the  
 CC invention. The sequences are useful for detecting the methylation state  
 CC of all Cpg dinucleotides in a sequence and therefore for analysing  
 CC associated diseases. By analysing cytosine methylations in the pretreated  
 CC DNA, genetic and/or epigenetic parameters for the diagnosis and therapy  
 CC of existing diseases or the predisposition to specific diseases can be  
 CC ascertained. The parameters may be compared to another set of genetic  
 CC and/or epigenetic parameters, the differences serving as basis for  
 CC diagnosis and/or prognosis events which are disadvantageous to patients.  
 CC The sequences of the invention are useful for the diagnosis and therapy  
 CC of HIV infection, neurodegenerative disorders, graft-versus-host disease,  
 CC aging, glomerular disease, Lewy body disease, arthritis,  
 CC arteriosclerosis, solid tumours and cancers.  
 CC  
 SQ Sequence 5935 BP; 1948 A; 61 C; 1039 G; 2887 T; 0 other;  
 Query Match 8.0%; Score 49.8; DB 22; Length 5935;  
 Best Local Similarity 47.3%; Pred. No. 0.048;  
 Matches 183; Conservative 0; Mismatches 202; Indels 2; Gaps 1;  
 QY 227 TATATGCTTATTTTATATGTTGCTCAAGATAGCTATTTATATGCAATATATAC 286  
 DB 2608 TTTAAGATTTTATTTAGTCTTTTGAAGGTGTTTATTTATTTATTAATAT 2667  
 QY 287 TTTTATGCTGTACGCGATGATTCATTACGTACAAGTTGTCATTATGACTA 346  
 DB 2668 TTTATAGTTAGATTTGTTATTTGTTT--TTATATTAATGCAATTTGTGAGGAATTAAT 2725

QY 347 CAAGGTTTTTAATTAATTAATGAGTAATTTATATATAATTAAGGTTGAAGAAATG 406  
 DB 2726 ATGAATTTTATTAATTAATGTTTATATTAATTAATTAATTTTATTAATAGTGT 2785  
 QY 407 TTGCTTCTCAACAACACTGTTGACTTTTACGTAGTTAGTTCAAGTAATTTATAG 466  
 DB 2786 TAGATTTTAAATTAATTTATTAATTTATTAATTAATTAATTAATTAATTTAG 2845  
 QY 467 GTATGTCATATATGATTAATTAAGAGAGTTTCAACTTTTATCCATTAATATAG 526  
 DB 2846 ATTTTAAATTAATTAATTAATTAATTAATTTTATTTAGTTTATTTATTTGTGG 2905  
 QY 527 CAAGTCTTCAGGTGGCAGATTAAATGCTTAATACAGGAATAGTATTTTATATA 586  
 DB 2906 GTTGATTTTGAATTTATTTATTTGTTTATTTATTAATGATGATAATTAATATCA 2965  
 QY 587 CGCTAATTCCTATGACATTTATTCAG 613  
 DB 2966 TTGATTAAGAGATGATTTTGTGAG 2992  
 RESULT 28  
 ABL34072  
 ID ABL34072 standard; DNA; 40862 BP.  
 XX  
 AC ABL34072;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Human immune system associated gene SEQ ID NO: 2045.  
 XX  
 KW Human; immune system disease; cytosine methylation; antiasthmatic;  
 KW antiarteriosclerotic; antianemic; cytostatic; neurotropic;  
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
 KW antirheumatic; antiarthritis; antidiabetic; antipsoriatic;  
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
 KW gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200200928-A2.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 02-JUL-2001; 2001WO-EP07537.  
 XX  
 PR 30-JUN-2000; 2000DE-1032529.  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX  
 PA (EPIG-) EPIGENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K;  
 XX  
 DR WPI; 2002-130909/17.  
 XX  
 PT Nucleic acid comprising fragment of chemically modified gene, useful  
 PT for diagnosis and treatment of diseases associated with abnormal  
 PT cytosine methylation -  
 XX  
 PS Claim 1; SEQ ID NO 2045; 32pp + Sequence Listing; German.  
 XX  
 CC The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
 CC diseases. The present sequence is a gene of the invention.  
 CC  
 SQ Sequence 40862 BP; 10964 A; 414 C; 8171 G; 21313 T; 0 other;

Query Match 8.0%; Score 49.8; DB 24; Length 40862;  
Best Local Similarity 45.2%; Pred. No. 0.072;  
Matches 183; Conservative 0; Mismatches 222; Indels 0; Gaps 0;

QY 108 TACTGACGACGAGGAGTGTGGGGCTTGTCTATTATTATTCAGGCGTAATGGGAGCTCATTTGA 167  
DB 16338 TATTGGGAATGTATAGTGTGGTTTAAATGTAGTGTGTAATTAATGTTAT 16397  
QY 168 AACATTTTCTACATATTTTACAGATCATTTCCGTTGAGGTTACAGACAGTCAGT 227  
DB 16398 CGATTAAATTTTATTTGTGTATGTATGTATTTAGATTGCGAGTTATTTATATATGA 16457  
QY 228 ATATGCTTTATTTATTTATTTGCTCAAGATAGTATTTTATGATGCCAATATTACT 287  
DB 16458 ATATTATTTTATTTATTTAGTGTGGTAATATAGAGATTTTGTAGATTTGTTTAT 16517  
QY 288 TTTTATTTGCTGTTACGGCATGATTTCTATACGTATCAAGTTGGTGCATTATGACTAC 347  
DB 16518 TGGGTTTTTTTTTTTAAAGATAGATTTATTTATTTATTTTAAAGTGTAT 16577  
QY 348 AAAGTTTAAATTTAAATGAGTAATTTATATATATAAAGGTTGAAGGAATGTT 407  
DB 16578 AATTAGTGTTTTATAGTATAGTATTAATTTGATATTTATTTAGAAATTTT 16637  
QY 408 TGCCTTCACAAACAACCTGTGACTTTTACGATTTAGTTCAAGTATTTGTTATAGG 467  
DB 16638 TATTATTTTGAAGAAATTTTATTTATTTATAGTATTTTATTTTTCGT 16697  
QY 468 TATTGTTCCATATATGATTTTAAAGAGAGTTTCAACTTTT 512  
DB 16698 TAGTTTTGGTAATTTATTTATTTTGTTTTATGATTTGTT 16742

RESULT 29  
ID AA201425 standard; DNA; 1038602 BP.  
XX AA201425;  
AC AA201425;  
XX  
DT 07-OCT-1999. (first entry)  
XX  
DE Complete genome sequence of Chlamydia trachomatis.  
XX  
KM Vaccine; eye disease; conventional trachoma; nongonococcal urethritis; epididymitis; cervicitis; salpingitis;  
KM nongonococcal urethritis; epididymitis; cervicitis; salpingitis;  
KM bartolinitis; pneumopathy; venereal lymphogranulomatosis; ss.  
XX Chlamydia trachomatis.  
OS  
XX  
PN WO928475-A2.  
PD 10-JUN-1999.  
XX  
XX 27-NOV-1998; 98WO-IB01939.  
PF  
XX  
XX 04-NOV-1998; 98US-0107077.  
PR 28-NOV-1997; 97FR-0015041.  
PR 17-DEC-1997; 97FR-0016034.  
XX  
XX (BEST ) GENSET.  
PA  
XX  
XX Griffais R;  
PI  
XX  
XX WPI; 1999-371125/31.  
DR  
XX  
XX  
XX Genome sequence of Chlamydia trachomatis  
XX  
XX Claim 1; Page 373-656; 1755pp; English.  
XX  
XX The present sequence represents the complete genome of Chlamydia  
CC trachomatis. Open reading frames (ORFs) of the genome encode

CC polypeptides AAY36754-Y37949. The polypeptides can be used as vaccines  
CC against Chlamydia trachomatis. Antisense and ribozyme sequences can also  
CC be used to control growth of the microorganism. Chlamydia trachomatis is  
CC responsible for a large number of diseases, e.g. eye diseases such as  
CC conventional trachoma, nongonococcal urethritis, epididymitis, and inclusion  
CC conjunctivitis; genital diseases such as nongonococcal urethritis;  
CC epididymitis, cervicitis, salpingitis, peritphalitis, bartolinitis;  
CC pneumopathy in breast feeding infants; and venereal  
CC lymphogranulomatosis. The polypeptides of the invention may be of use in  
CC treating these diseases.  
XX  
SQ Sequence 1038602 BP; 304265 A; 214645 C; 214259 G; 305001 T; 432 other;

Query Match 8.0%; Score 49.8; DB 20; Length 1038602;  
Best Local Similarity 47.3%; Pred. No. 0.14;  
Matches 184; Conservative 0; Mismatches 202; Indels 3; Gaps 1;

QY 19 AAAACAGAGAAAGCAACCCGAAAGCAGTCAAGAAAGCTGTTCTGAAGGAGTGTCCCT 78  
DB 364123 AAAACAGAAAGGCGACCCCAAGCGTTCGGGACGTAAGAAAAGGGGAGTATGCC 364182  
QY 79 AATCAGAAAGAGTTTCTAAGCATGACTGACGAGGATGCTGGGCTTGTCTATT 138  
DB 364183 AATCTCAAGATTTTCTCTGATTAAGTTCATTTCTCAATGTTCTTAAGCTTTTCT 364242  
QY 139 TATTCAGGCGTATGAGGAGTCAATTTTGAACAATTTTCTACTATATTTTACAGATCA 198  
DB 364243 TTGGCTCATTTTTCCTGCAACCACTTAGAGAGTTTATGATCTATTTTAAACAGCT 364302  
QY 199 TTTGGTTTGAAGTTACAGACAGTCAATATGCTTTATTTATTTATTTATGTTGCTCAAG 258  
DB 364303 CCCCAATCATGATCGGACCTTAGCATCTATTTTAAATAATGCTTATTC- 364359  
QY 259 ATAGCTATTTTATGATGCAATATTTACTTTTATTTGCTGTACGCAATGATTTCAAT 318  
DB 364360 ATTCTGACAGTCTCTCTCTTCTTTGTTAGAGCGGTGGAATGATGATTTGATAGGA 364419  
QY 319 CGGTACAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 378  
DB 364420 TTTCTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 364479  
QY 379 AATATATATAAAGGTTGAAGGATGTT 407  
DB 364480 AATCCATTGACAACTGAACAAAATTT 364508

RESULT 30.  
ID AAS46788 standard; DNA; 61020 BP.  
XX AAS46788;  
AC AAS46788;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Tumour suppressor gene derived chemically modified sequence #514.  
XX  
XX  
XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic;  
XX cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;  
XX cytosine methylation; de.  
XX  
XX Homo sapiens.  
OS  
XX  
XX  
XX WO200168912-A2.  
PN  
XX  
XX 20-SEP-2001.  
PD  
XX  
XX  
XX 15-MAR-2001; 2001WO-EP02955.  
PF  
XX  
XX 15-MAR-2000; 2000DE-1013847.  
PR 06-APR-2000; 2000DE-1019058.  
PR 07-APR-2000; 2000DE-1019173.  
PR 30-JUN-2000; 2000DE-1032529.  
PR 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPIGENOMICS AG.  
 PA Olek A, Piepenbrock C, Berlin K;  
 PI WPI; 2001-602752/68.  
 XX  
 XX  
 XX Fragments of chemically modified genes associated with tumour suppressor  
 PT genes and oncogenes, useful in designing primers and probes for  
 PT analysing diseases associated with cytosine methylation state e.g.  
 PT cancer  
 PS Claim 1; SEQ ID NO 514; 27bp; English.  
 XX  
 XX The invention relates to a nucleic acid comprising a sequence of 18  
 CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with  
 CC bisulphite, of genes associated with tumour suppression and  
 CC oncogenes having a sequence taken from 536 (actually 533 since  
 CC numbers 408, 458 and 500 are missing from the sequence listing) sequences  
 CC (Ss) and sequences complementary to (Ss). The nucleic acid may be a  
 CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may  
 CC form part of a set of probes for detecting the cytosine methylation state  
 CC and/or single nucleotide polymorphisms and also to be used in an  
 CC array for analysing diseases associated with CpG dinucleotides e.g.  
 CC cancers and tumours. The probes can also be used in a method for  
 CC ascertaining genetic and/or epigenetic parameters for the diagnosis  
 CC and/or therapy of existing diseases or the predisposition to specific  
 CC diseases, by analysing cytosine methylations. The parameters may be  
 CC compared to another set of genetic and/or epigenetic parameters, the  
 CC differences serving as basis for diagnosis and/or prognosis events which  
 CC are disadvantageous to patients. The present sequence is one of the  
 CC 533 genomic sequences derived from tumour suppressor genes and  
 CC oncogenes. Sequences with even numbered Seq ID numbers are the  
 CC complementary sequence of the corresponding odd numbered sequence (e.g.  
 CC ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence  
 CC is missing).  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 61020 BP; 17884 A; 634 C; 12537 G; 29965 T; 0 other;

Query Match 7.9%; Score 49.2; DB 22; Length 61020;  
 Best Local Similarity 45.5%; Pred. No. 0.11;  
 Matches 174; Conservative 0; Mismatches 208; Indels 0; Gaps 0;

OY 173 TTTCTACTATATTTTACAGATCATTCGGTTGAGGTTACAGACAGTCAGTATATG 232  
 |||||  
 DB 57697 TTTTCTACTATATTTTACAGATCATTCGGTTGAGGTTACAGATATAGGTTTATAGGTTAGT 57756  
 |||||  
 OY 233 CTATATATTTATTTGCTCTCAAGATAGCTATTTATATGATGCAATATTTACTTTTA 292  
 |||||  
 DB 57757 TTGATTTTATTTGATTTTACGAGAAATACGTTTATATATATTTTATTTTATTTTA 57816  
 |||||  
 OY 293 TTGCTGTTACGGCATGATTTTATTCATTCAGTCAAGTTGGTCATTAAGACTACAAAG 352  
 |||||  
 DB 57817 TCGCGGATGAAATTTAGGTTATTTATTTTGTGTTTTTTTTTTTTTTTTTTTTT 57876  
 |||||  
 OY 353 TTTTAAATTTAAATGAGTAATTTAATTAATAAAGGGTGAAGAAATGTTGCTT 412  
 |||||  
 DB 57877 TTTTAAATTTAAATGAGTAATTTAATTAATAAAGGGTGAAGAAATGTTGCTT 57936  
 |||||  
 OY 413 CTCACAAACACTTGTGCACTTTTACGATTTAGTTCAAGTAATGTTATAGGTAATG 472  
 |||||  
 DB 57937 TTTTAAATTTAAATGAGTAATTTAATTAATAAAGGGTGAAGAAATGTTGCTT 57996  
 |||||  
 OY 473 TTCCATATATGATTTAATAAGAGAGTTTCAAACTTTTACCATTAATATATGCAAGTC 532  
 |||||  
 DB 57997 TTGATATATGATTTTGAAGAAATTAATTTAGATATAGGTAATTTTATAGGTAAGA 58056  
 |||||  
 OY 533 CTTCAGGTGTGCGACATTTATAT 554  
 |||||

DB 58057 TTAAGTTTGAATGATTTT 58078

RESULT 31  
 ID ABL34021  
 ID ABL34021 standard; DNA; 5542 BP.  
 XX  
 AC ABL34021;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 XX Human immune system associated gene SEQ ID NO: 1994.  
 DE  
 XX Human; immune system disease; cytosine methylation; antiasthmatic;  
 KW antiasthmatic; antiasthmatic; antiasthmatic; antiasthmatic;  
 KW antiproliferative; anti-HIV; anticonvulsant; ophthalmological;  
 KW antineoplastic; antineoplastic; antineoplastic; antineoplastic;  
 KW antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;  
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
 KW gene; ds.  
 XX  
 XX Homo sapiens.  
 XX  
 XX WO200200928-A2.  
 PN  
 XX  
 PD 03-JAN-2002.  
 XX  
 PD 02-JUL-2001; 2001WO-EP07537.  
 PF  
 XX 30-JUN-2000; 2000DE-1032529.  
 PR 01-SEP-2000; 2000DE-1043826.  
 PR  
 XX (EPIC-) EPIGENOMICS AG.  
 PA Olek A, Piepenbrock C, Berlin K;  
 PI WPI; 2002-130909/17.  
 DR  
 XX  
 XX Nucleic acid comprising fragment of chemically modified gene, useful  
 PT for diagnosis and treatment of diseases associated with abnormal  
 PT cytosine methylation  
 PT  
 XX  
 PS Claim 1; SEQ ID NO 1994; 32pp + Sequence Listing; German.

XX The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
 CC diseases. The present sequence is a gene of the invention.  
 CC  
 SQ Sequence 5542 BP; 1715 A; 55 C; 1041 G; 2731 T; 0 other;

Query Match 7.8%; Score 48.8; DB 24; Length 5542;  
 Best Local Similarity 54.8%; Pred. No. 0.081;  
 Matches 160; Conservative 0; Mismatches 127; Indels 5; Gaps 3;

OY 226 GTATATGCTTATTTATTTATTTGCTCTCAAGATAGCTATTTATATGATCCAAATTA 285  
 |||||  
 DB 2468 GTATATGCTTATTTATTTATTTGCTCTCAAGATAGCTATTTATATGATCCAAATTA 2524  
 |||||  
 OY 286 CTTTAAATTTGCTGTTACGCAATGATTTTATTCATTCAGTCAAGTTGGTCATTTAGACT 345  
 |||||  
 DB 2525 ATTTTAAATTTGCTGTTGTTGTTTATTTTATTTGCTGTTATTAATATTTTAA-T 2583  
 |||||  
 OY 346 ACAAGGTTTAAATTTAATGAGTAATTTAATTAATAAAGGGTGAAGAAATG 405  
 |||||  
 DB 2584 ATTAAGAAATTTTAAATCGTATTTTGAATATTTTGTGTTATTTTGTGTAATAT- 2642  
 |||||  
 OY 406 TTGCTTCTCAACAAACACTTGTGCACTTTTACGATTTAGTATTTCAAGTAATTTGTATA 465  
 |||||

Db 2643 TTGTTTGTAAATAATTTTAAAGATTAATATAGATTAATAAGATTTTAA 2702  
 Qy 466 GGTATGTCATATGATTAATAAGAGAGTTTCAACTTTACAT 517  
 Db 2703 GTGATTTATATTAATGATTTTATATGAAATTTAAAGTTAAATTTAT 2754  
 RESULT 32  
 AAS63347  
 ID AAS63347 standard; DNA; 5857 BP.  
 AC AAS63347;  
 XX  
 DT 29-JAN-2002 (first entry)  
 XX  
 DE Chemically pretreated metabolism associated gene #42.  
 XX  
 KM Human; cytostatic; anti-tumour; metabolism; metabolic disease; liver;  
 KM solid tumour; cancer; cytosine methylation; epigenetic; eye; kidney;  
 KM single nucleotide polymorphism detection; SNP; stool; urine; lung;  
 KM cerebral-spinal fluid; intestine; brain; heart; prostate; breast;  
 KM DUSP2; ERHX2; QDPR; SGRH; SHMT2; SLC7A2; SLC7A4; TYMS; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200176451-A2.  
 XX  
 PD 18-OCT-2001.  
 XX  
 PE 06-APR-2001; 2001WO-EP04016.  
 XX  
 PR 06-APR-2000; 2000DE-1019058.  
 PR 07-APR-2000; 2000DE-1019173.  
 PR 30-JUN-2000; 2000DE-1032529.  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX  
 PA (EPIG-) EPIGENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K;  
 XX  
 DR WPI; 2002-010834/01.  
 XX  
 PT New nucleic acid, useful for diagnosis and therapy of metabolic  
 PT disease, solid tumour and cancers, comprises segment of chemically  
 PT modified genomic sequences of genes associated with metabolism  
 PS  
 PS Claim 1; Page 113-114; 143bp; English.  
 CC The invention relates to a nucleic acid (I) comprising a sequence at  
 CC least 18 bases of a segment of the chemically pretreated DNA of genes  
 CC associated with metabolism such as DUSP2 (NM\_004418), ERHX2 (NM\_001979),  
 CC QDPR (NM\_000320), SGRH (NM\_000199), SHMT2 (NM\_005412), SLC7A2  
 CC (NM\_003046), SLC7A4 (NM\_004173) and TYMS (NM\_001071) (all  
 CC undefined). (I) are useful for diagnosis and therapy of metabolic  
 CC disease, solid tumours and cancers; as primer oligonucleotides for the  
 CC amplification of DNA sequences, for detecting the cytosine methylation  
 CC state and/or single nucleotide polymorphisms (SNPs) in a chemically  
 CC treated DNA of genes associated with metabolism. An array of (I) is  
 CC useful for ascertaining genetic and/or epigenetic parameters for the  
 CC diagnosis and/or therapy of existing diseases or the predisposition to  
 CC specific diseases by analysing cytosine methylations. The method involves  
 CC chemically treating genomic DNA sample by a solution of bisulphite,  
 CC hydrogen sulphite or disulphite such that cytosine bases which are  
 CC unmethylated at the 5th-position are converted to uracil or another base  
 CC which is dissimilar to cytosine in terms of hybridisation behaviour and  
 CC amplifying fragments of the chemically pretreated genomic DNA. The  
 CC genomic DNA is from cells or cellular components which contain DNA,  
 CC sources of DNA comprising, for e.g. cell lines, biopsies, blood, sputum,  
 CC stool, urine, cerebral-spinal fluid, tissue embedded in paraffin such as  
 CC tissue from eye, intestine, kidney, brain, heart, prostate, lung, breast  
 CC or liver, histologic object slides and their combinations. Genetic  
 CC parameters are mutations, in particular insertions, deletions, point

CC mutations, insertions and polymorphisms of genes associated with  
 CC metabolism and sequences are further required for their regulation.  
 CC Epigenetic parameters are in particular cytosine methylations and  
 CC further chemical modifications of DNA bases of genes associated with  
 CC metabolism. Further epigenetic parameters include for e.g. the  
 CC acetylation of histones which correlates with DNA methylation.  
 CC AAS63306-AAS63373 represent chemically pretreated metabolism associated  
 CC genes, and related primers of the invention.  
 XX  
 SQ Sequence 5857 BP; 1535 A; 131 C; 1530 G; 2623 T; 38 other;  
 Query Match 7.8%; Score 48.8; DB 24; Length 5857;  
 Best Local Similarity 49.6%; Pred. No. 0.082;  
 Matches 122; Conservative 0; Mismatches 124; Indels 0; Gaps 0;  
 Qy 163 TTGGAACATTTCTACTATTTTACAGATCATTTGGTTAGGTTACAGCAG 222  
 Db 3551 TTGGAGGAATGTTAGTTTNTTTTATTTATTTTGGGGTGGCGGTTGGTTG 3610  
 Qy 223 TCAGTATATGCTTATTTATTTATTTATTTGCTCAAGATAGCTATTTATTTGATGCCAATA 282  
 Db 3611 TTATTTAAGATATTTTATTTAGTTAGGTTGAAGGAGATGAAATGTTATTTGGGAATA 3670  
 Qy 283 TTACTTTTATTTCTGTTACGCGATGATTTCATTCGTGTACAACTTGTCATTATYGG 342  
 Db 3671 TTGTTTNTTTTGTGTAGTTGATTTGAATGTTAAGTTATGTTTCGTTAGGGTTGT 3730  
 Qy 343 ACTACAAAGTTTATTTATTTAATGATGATGATTTATTTATTTAAGGTTGAAGA 402  
 Db 3731 TTGTTAGTTTATTTTATTTAATGAGGAGATGATGAGATGATGAGATGAGATG 3790  
 Qy 403 ATGTTT 408  
 Db 3791 TTATTT 3796  
 RESULT 33  
 ABL32610  
 ID ABL32610 standard; DNA; 18855 BP.  
 XX  
 AC ABL32610;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Human immune system associated gene SEQ ID NO: 583.  
 XX  
 KM Human; immune system disease; cytosine methylation; antiaesthetic;  
 KM antiarteriosclerotic; antiamaemic; cytostatic; nootropic;  
 KM neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
 KM antihypertensive; antiaortic; antidiabetic; antiparasitic;  
 KM antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;  
 KM acute myeloid leukemia; Alzheimer's disease; AIDS; epilepsy;  
 KM neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
 KM gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200200928-A2.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 02-JUL-2001; 2001WO-EP07537.  
 XX  
 PR 30-JUN-2000; 2000DE-1032529.  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX  
 PA (EPIG-) EPIGENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K;  
 XX  
 DR WPI; 2002-130909/17.  
 XX  
 PT Nucleic acid comprising fragment of chemically modified gene, useful

PT for diagnosis and treatment of diseases associated with abnormal  
PT cytosine methylation -

PS Claim 1; SEQ ID NO 583; 32pp + Sequence Listing; German  
vv

The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines. The sequences can be used in the diagnosis and treatment of immune system disorders, including eye diseases such as retinopathy, neovascular glaucoma and macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis, rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel diseases. The present sequence is a gene of the invention.

Sequence 18855 BP; 5520 A; 178 C; 3950 G; 9207 T; 0 other;

Query Match	7.8%	Score	48.6	DB	24	Length	18855
Best Local Similarity	46.9%	Pred	No.0.12				
Matches	187	Conservative	0	Mismatches	209	Indels	3
						Gaps	1

QY	125	TGGGGCTGCTAATTTATTCAGGCGCTAATGGAGCGCTATTGTGAAACAATTTCTACTA	184
Db	10618	TAGGAGTTAATTTTTTATTATTATTTTAAAGGGATATATTTTTTTTAGTAGTATATTTA	10677
QY	185	TTTTTAACAATCAATTCGGTTGAGGTACAGACAGTCAGATATGCTTATTTATTT	244
Db	10678	GTTTTTTAAAAAGTTTATGTTTTTTGTTTTTAAATATATGGGGTTTTTTTTTATTTGTT	10733
QY	245	ATGTGTCAAGATAGCTATTTTATTGATGCAATATTACTTTTATTTGCTGTACGG	304
Db	10738	ATTATATAGATGAATTAATTTATTTTTTTTATTTAATTAGTGTTAGTTGTTAATT	10797
QY	305	CATGGATTTCAATTCAGTGTACAAGTTG--GTGCATATGAGCACTCAAAAGTTTTAAAT	361
Db	10798	TTTGCGTAATGTTTATTTTATTTTAAATTTATAGTGTTGAAATTTAAATATTTAAAG	10855
QY	362	TTAAATGAGTAAATTTAATATATAATTAATAAGGTTGAAGAAGATGTTGCTTCCAA	421
Db	10858	TTATTTGTTGTAATAATAATGAATTAATGTCGAAGAGGTTAGTTTTTTGTTTAGAA	10912
QY	422	CACTTGTTCCACTTTTACGTAGTTAGTTCAAGTAATGTTATAGGTAATGTTCCATATA	481
Db	10918	ATAAAGTGTAGTGTATATTTTATTTAGGTAGGTTTAGAATTTAAATATGATTTTTCG	10977
QY	482	TGATTTATAAAGGAGATTTTCAACCTTTTACACTAT	520
Db	10978	TTTTTTTTATAGTTTTTTTGTTTTTTTTTATATATTTT	11016

RESULT 34  
ABL92243  
ID ABL92243 standard; DNA; 11787 BP.

AC	ABL92243;
XX	
DT	01-JUL-2002 (first entry)
..	

Chemically treated DNA repair gene fragment complementary to #26.

KM DNA repair: cytosine methylation; PMS2L1; PMS2L2; PMS2L3;  
KM PMS2; L4; PMS2L5; PMS2L6; MGMT; MSH2; NUDT1; TDG; INPPL1; RFC4;  
KM DITTL1; FANCB; XRCC8; ataxia telangiectasia; aging; Bloom's syndrome;  
KM Cockayne syndrome; Nijmegen breakage syndrome; Werner syndrome;  
KM immunodeficiency; trichothiodystrophy; Fanconi's anaemia; solid tumour  
KM cancer; ds.

OS Unidentified.

PN WO200181622-A2.

PD 01-NOV-2001.

PF 06-APR-2001; 2001WO-EP03972.

XX 06-APR-2000; 2000DE-1019058  
PR 07-APR-2000; 2000DE-1019173  
PR 30-JUN-2000; 2000DE-1032529  
PR 01-SEP-2000; 2000DE-1043826

PA (EPiG-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K, xv

DR WPI; 2002-034446/04

PT New nucleic acid derived from genes associated with DNA repair, useful  
PT for diagnosis, e.g. of ataxia telangiectasia, by determination of  
PT cytosine methylation -

PS Claim 1; SEQ ID NO 52; 25bp + sequence listing; English.  
....

The invention relates to nucleic acids containing a sequence of at least 18 nucleotides of chemically treated DNA of genes associated with DNA repair, and their complements. The invention also relates to nucleic acids comprising at least 18 base pairs of the chemically pretreated DNA of genes associated with DNA repair selected from PMS211, PMS2112, PMS212, PMS213, PMS2, L4, PMS215, PMS216, MGMT, MSH2, NDDT1, TIG, INPPL1, R44, DDT11L1, FANCB, or XRCC8. Nucleic acids of the invention and related oligomers, are useful for diagnosis of diseases associated with gene repair, specifically ataxia telangiectasia, aging, Bloom's syndrome, Cockayne syndrome, trichothiodystrophy, Fanconi's anaemia, solid tumours and cancer, particularly by determining status of cytosine methylation and/or by detecting single-nucleotide polymorphisms. Determination of individual methylation patterns may allow development of individualised therapies. The sequences given in records ABL92197-ABL92315 represent chemically pre-treated DNA fragments from genes associated with DNA repair and their complements.

Note: The sequence data for this patent is not represented in the specification, but is based on sequence information supplied by the European Patent Office.

SQ Sequence 11787 BP; 3541 A; 134 C; 2217 G; 5895 T; 0 other;

Query Match	7.8%;	Score 48.4;	DB 24;	Length 11787;
Best Local Similarity	46.7%;	Pred. No. 0.12;		
Matches 154;	Conservative 0;	Mismatches 176;	Indels 0;	Gaps 0

Qy	159	TCATTTGAAACAATTTTCTACATATATTTTACGAATCACTTGGTTGAGGTACAGC	218
Db	4043	TTATTTAGATTATTTTGTGTTAGTAAGTAATAATTTAGTAGTTTATATATGATTTT	4102
Qy	219	ACAGTCAGTATAGCTTATTTATTTATTTAGTTCCTCAAGAGATAGCTATTTATATGATGCC	278
Db	4103	TTTGTGTTTTTGACGTGTTTTTTTAAAGATGCGTTTACTGTGTTTTTTTGGTTCGT	4166
Qy	279	AATATTACTTTTATTCCTGTATACGCAATGATTTCAACGTACAAAGTTGGTCATT	338
Db	4163	ATTTTTTTTTTATTTATTTTGAAGATATGTTAAGGTGATTTATTTTGTGAATGCTT	4222
Qy	339	ATGCACACAAAGGTTTTTAATTAATTAATGAGTAAATTTATATTAATTAAGGGTTCAA	398
Db	4223	TTTTTTATTTTGTGTTTTATTTATTAATTAATTTATTTTAAATAATTAATTAATTTTA	4282
Qy	399	AGGAATGTTGCTCTCAACAAACACTGTTCGACTTTTACGTAGTTTAGTTCAAGTAT	458
Db	4283	AATATAGCTTATTTTAAAGTGTATTTGTGATTTTAAAGTTGTGTTTTTTTATATGAA	4342
Qy	459	TGTTATAGGTAATGTCATATATATGATTAAT	488
Db	4343	TTTATGTTTATAGAGTATATGATTTATAT	4372

RESULT 35  
ABN80017  
ID ABN80017 standard; DNA; 19345 BP



XX AC ABRN0017;  
 XX DT 15-JUL-2002 (first entry)  
 XX DE Human chemically modified disease associated gene SEQ ID NO 34.  
 XX KW Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;  
 KW heart disease; epilepsy; histone deacetylation; muscular dystrophy;  
 KW dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;  
 KW antidiabetic; cytosolic; anticonvulsant; ds.  
 OS Homo sapiens.  
 OS Synthetic.  
 XX MO200200927-A2.  
 XX PD 03-JAN-2002.  
 XX PF 02-JUL-2001; 2001WO-EP07536.  
 XX PR 30-JUN-2000; 2000DE-1032529.  
 XX PR 01-SEP-2000; 2000DE-1043826.  
 XX PA (EPIG-) EPIGENOMICS AG.  
 XX PI Olek A, Piepenbrock C, Berlin K;  
 XX WPI; 2002-130908/17.  
 XX DR  
 XX PT Novel nucleic acid useful for diagnosis and therapy of diseases  
 PT associated with development genes such as diabetes, comprises a  
 PT sequence of a segment of chemically pretreated DNA of genes associated  
 PT with development  
 XX  
 XX Claim 1; SEQ ID NO 34; 27bp; English.  
 XX  
 CC The invention relates to a nucleic acid (I) comprising a sequence at  
 CC least 18 bases in length of a segment of chemically pretreated DNA (II)  
 CC of genes associated with development selected from 87 genes listed in  
 CC the specification such as ACCFN, ADFN, or AFDI and comprising one of 350  
 CC sequences (ABN79984-ABN80333) or their complements. The invention is  
 CC useful for the diagnosis or therapy of diseases associated with  
 CC development genes, in particular disease related to homeobox containing  
 CC genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes  
 CC associated with congenital heart disease, epilepsy, diseases related to  
 CC histone deacetylation, Currarino syndrome, diseases related with the  
 CC development of the brain and limb girdle muscular dystrophy and dwarfism.  
 CC Oligomers specific to each of the genes are useful for detecting the  
 CC methylation state of all CpG dinucleotides within the 350 sequences or  
 CC (II) and their complementary sequences, as primer oligonucleotides for  
 CC the amplification of the 350 sequences, (II) and/or their complements and  
 CC as oligomer probes for detecting the cytosine methylation state and/or  
 CC single nucleotide polymorphisms (SNPs).  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification but is based on sequence information supplied to Derwent by  
 CC the European Patent Office.  
 XX  
 SQ Sequence 19345 BP; 5413 A; 464 C; 3983 G; 9485 T; 0 other;  
 Query Match 7.8%; Score 48.4; DB 24; Length 19345;  
 Best Local Similarity 45.5%; Pred. No. 0.13;  
 Matches 212; Conservative 0; Mismatches 251; Indels 3; Gaps 1;  
 QY 139 TATTAGGCGTATGAGACGATCTTTTGAACAATTTCTACTATATTTTACAGATCA 198  
 DB 12502 TTTTATATAGTAGAGAAAAGTTTAAAGTTATATAGTATAAGTTATGAAAATTT 12561  
 QY 139 TTTGCGTTGAGGTACAGACAGTCAGATATGCTTATATTTATTTAGTGTCTAAGAG 258  
 DB 12562 TGTGCGATTTAGGTTTATTTAGGTTTACAGATGTTGTTTGTGTTGGGAAATGT 12621  
 QY 259 ATAGCTATTTATTTAGATGCATATTTACTTTTATTTGCTGTATACGCAATGATTTCAATTA 318

DB 12622 TTATACGTAATGTTAAAGATGAAATAGATTTTTTTTTTTTTTTCGAAGGAATGAGTT 12681  
 QY 319 CGGTACAGATGGTGCAATATGACATCAACAAGTTTAAATTAATGAGTAAATTT 378  
 DB 12682 ATTATATTAAGAGATGATTTGTGTAAATATATATTTATTTATTTATTTAGTA 12741  
 QY 379 AATATTA---TAAAGGTTGAAGAAGATGTTTCTCTCAACAACACTGTTCGACTT 435  
 DB 12742 TGTATGCTTTTACGGTTTATATTTATTTATTTAATGTAATAATTTAGTTTACG 12801  
 QY 436 TTACGATGTTTACGTTCAAGTAAATTTTAAAGTATGTTCCATATATGATTTAAAGGA 495  
 DB 12802 AGATTATTTTGTTTTAAATTTTATTTTGTGTTTATATATTTTGTGTTT 12861  
 QY 496 GAGTTTCAAACTTTTACATTTATTTATTTATGCAAGTCTCAGGTGGCAGATTAATG 555  
 DB 12862 AAAAAATTTAGTATATTAAGGTTGTATGGAATTTAGTTTAAATTTATTTATTAAGTA 12921  
 QY 556 CTTAATACAGGAATAGTACTGTTTATTAATACGTAATTCCTATGA 601  
 DB 12922 AATTATTTTAAAGATATTTGATATATGATATATTTATTTTAA 12967  
 RESULT 36  
 ABRK40034  
 ID ABRK40034 standard; DNA; 12138 BP.  
 XX AC ABRK40034;  
 XX DT 21-MAY-2002 (first entry)  
 XX DE Human chemically pretreated gene sequence #58 strand 2.  
 XX KW Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;  
 KW cytosolic; ALDH6; CYP11A; CYP11B; CYP3A3; DPYD; EPHX2; OCIN; TXNRD1;  
 KW UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.  
 OS Homo sapiens.  
 XX MO200202806-A2.  
 XX PD 10-JAN-2002.  
 XX PF 29-JUN-2001; 2001WO-EP07470.  
 XX PR 30-JUN-2000; 2000DE-1032529.  
 XX PR 01-SEP-2000; 2000DE-1043826.  
 XX PA (EPIG-) EPIGENOMICS AG.  
 XX PI Olek A, Piepenbrock C, Berlin K;  
 XX WPI; 2002-154757/20.  
 XX DR  
 XX PT New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,  
 PT useful for detecting cytosine methylation state of genes associated  
 PT with pharmacogenomics and for therapy of diseases e.g. cancer  
 XX  
 XX Claim 1; SEQ ID NO 116; 24bp; English.  
 XX  
 CC The invention relates to a nucleic acid comprising a sequence at  
 CC least 18 bases in length of a segment of the chemically pretreated DNA  
 CC of genes associated with pharmacogenomics according to one of the  
 CC sequences of the genes ALDH6 (NM\_000693), CYP11A (NM\_000781), CYP11B1  
 CC (NM\_000497), CYP3A3 (NM\_000776 and NM\_017460), DPYD (NM\_000110), EPHX2  
 CC (NM\_001979), OCIN (NM\_002538), TXNRD1 (NM\_003330), UGT8 (NM\_003360),  
 CC MRP (NM\_004996, NM\_019900, NM\_019901, NM\_019902, NM\_019862, NM\_019898,  
 CC NM\_019859) and their complementary sequences, or a sequence (SI) chosen  
 CC from 87 sequences and their complements. The chemical pretreatment  
 CC is bisulphite treatment to convert cytosines (but not methyl-cytosines)  
 CC into uracils. Also included are an oligomer (II) in particular an  
 CC oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in



OS Homo sapiens.  
 XX WO200200928-A2.  
 XX  
 XX 03-JAN-2002.  
 XX  
 XX 02-JUL-2001; 2001WO-EP07537.  
 XX  
 XX 30-JUN-2000; 2000DE-1032529.  
 XX  
 XX 01-SEP-2000; 2000DE-1043826.  
 XX  
 XX (EPIG-) EPIGENOMICS AG.  
 XX  
 XX Olek A, Piepenbrock C, Berlin K;  
 XX  
 XX WPI; 2002-130909/17.  
 XX  
 XX Nucleic acid comprising fragment of chemically modified gene, useful  
 PT for diagnosis and treatment of diseases associated with abnormal  
 PT cytosine methylation -  
 XX  
 XX Claim 1; SEQ ID NO 1893; 32bp + Sequence Listing; German.  
 XX  
 XX The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
 CC diseases. The present sequence is a gene of the invention.  
 CC  
 XX Sequence 12590 BP; 2516 A; 310 C; 3233 G; 6531 T; 0 other;  
 SQ  
 Query Match 7.7%; Score 48.2; DB 24; Length 12590;  
 Best Local Similarity 45.0%; Pred. No. 0.13;  
 Matches 220; Conservative 0; Mismatches 268; Indels 1; Gaps 1;  
 Oy 80 AATCAGAGAGGTTCTAAGACATGACTACGACAGGAGATGCTGGGCTTCTATTT 139  
 Db 752 ATTTAGTGTAAATTAATTAAGAAATGATTTGTTTATGTTATGAGAGCGAATTGGA 811  
 Oy 140 ATTCAGCGGTAAATGAGCGTCAATTTGAACAATTTCTACTATATTTTACAGATCAT 199  
 Db 812 AATTAGTGTAAATTTGGAATAAATTAAGAGAAATTTGTTTGTATTTTATGATTTTGG 871  
 Oy 200 TTCGGTTGAGGTTACAGACAGTCAATATGCTTTAT-TTATTTATGTTCTCAAGAG 258  
 Db 872 GATTGTTTGAATTTTATTTGATTTGTTATTAATTTTATTTTAAATGTTAGCGTTTGG 931  
 Oy 259 ATTAGCTATTTATGATGCAATATTTACTTTTATTTGCTGTTAGCGCATGATTTTCATTA 318  
 Db 932 GAATTTTATTTTGTATTTATGTTATTAATTTTTCGTGTGTGTTAAATTTTGTGTTG 991  
 Oy 319 CGTGAAGAGTTGTCATTTATGACTCAAAAGTTTAAATTTAAATGAGATTAATTT 378  
 Db 992 TTTTATTTTGAAGAGATTAAGTTAGTTATGTTTATTTTAAATTTTATTAATTA 1051  
 Oy 379 AATATATATAAAGGTTGAAGAAGATGTTGCTTCTCAACAACACTGTTGCACTTTTA 438  
 Db 1052 TTTGTAATAAATTTTATTTTAAATAGATTAATTTATGTTTGAAGAAGATTTTGT 1111  
 Oy 439 CGTAGTTAGTTCAAGTATTTAGTTATGTTTCAATATTAATTAATAAAGAGAG 498  
 Db 1112 GGGGGTTTGTATTAATTAAGTTTGAATGCTTAATTTTGTGTTTAAATGAATTA 1171  
 Oy 499 TTTTCAACCTTTTACATTAATTAATTAATCAAGTCCCTGAGGTGAGCAATTAATGCTT 558  
 Db 1172 TTTATTAAGTTTGTGTTTATTTTATTTTGAATATTTGTAATGTTAATGTTTGA 1231  
 Oy 559 AATACAGA 567  
 Db 1232 TAGGTAGA 1240

RESULT 39  
 ABL32977  
 ID ABL32977 standard; DNA; 18512 BP.  
 XX  
 XX ABL32977;  
 XX  
 XX 26-MAR-2002 (first entry)  
 XX  
 XX Human immune system associated gene SEQ ID NO: 950.  
 XX  
 XX Human; immune system disease; cytosine methylation; antiasthmatic;  
 XX antiarteriosclerotic; antihaemic; cytostatic; nocotropic;  
 XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
 XX antirheumatic; antiarthritic; antidiabetic; antiporiatic;  
 XX antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
 XX acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
 XX neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
 XX gene; ds.  
 XX  
 XX Homo sapiens.  
 XX  
 XX WO200200928-A2.  
 XX  
 XX 03-JAN-2002.  
 XX  
 XX 02-JUL-2001; 2001WO-EP07537.  
 XX  
 XX 30-JUN-2000; 2000DE-1032529.  
 XX  
 XX 01-SEP-2000; 2000DE-1043826.  
 XX  
 XX (EPIG-) EPIGENOMICS AG.  
 XX  
 XX Olek A, Piepenbrock C, Berlin K;  
 XX  
 XX WPI; 2002-130909/17.  
 XX  
 XX Nucleic acid comprising fragment of chemically modified gene, useful  
 PT for diagnosis and treatment of diseases associated with abnormal  
 PT cytosine methylation -  
 XX  
 XX Claim 1; SEQ ID NO 950; 32bp + Sequence Listing; German.  
 XX  
 XX The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
 CC diseases. The present sequence is a gene of the invention.  
 CC  
 XX Sequence 18512 BP; 5520 A; 92 C; 3489 G; 9411 T; 0 other;  
 SQ  
 Query Match 7.7%; Score 48.2; DB 24; Length 18512;  
 Best Local Similarity 49.5%; Pred. No. 0.14;  
 Matches 152; Conservative 0; Mismatches 153; Indels 2; Gaps 1;  
 Oy 168 AACAAATTTCTACTATATTTTACAGATCAATTTGAGTTAGCTTACAGACACTCAGT 227  
 Db 4706 AATATTATAGTAATTTAGTTGAATGATTAATTTGTTTGAAGATATATAAATTTTGT 4765  
 Oy 228 ATATGCTTATTTATTTATGTTGCTCAAGATAGCTATTTTATTTAGCCAAATTTACT 287  
 Db 4766 TTTTATATTTGTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTATTT 4825  
 Oy 288 TTTTATGCTGTACGCGATGATTTCACTAGCTGTAACAATGTTGATGATGACTAC 347  
 Db 4826 GTGATTTGTTATTAATAGAAATGAAATGCTTTTATTAATGTTTATTTATTTATG-1GG 4883  
 Oy 348 AAAGTTTAAATTTAAATGAGTAATTTAATTAATTAATAAAGGTTGAAGAAGATGTT 407

Db	4884	GAAAGAAAAAGAGATATATAAATAATATAATATGGTTTTTATATATATTA	4943
Qy	408	TGCTTCACAAACACTTGTTCAGCTTTAGTATGATCAAGTAATGTATAG	467
Db	4944	TGTAGTTATTTTATATGAGAGTTTTTATTTTATATATATTTTAAGTAATGTTAATG	5003
Qy	468	TATGTGT	474
Db	5004	TTTTTTT	5010

RESULT 40  
ABN85766/c  
ID ABN85766 standard; CDNA; 82952 BP.  
vv

DT 21-OCT-2002 (first entry)

DE Arabidopsis yellow stripe1-like 3 encoding CDNA SEQ ID NO. 7.

KW Maize; transgenic; plant; yellow stripe1-like; ysl; Arabidopsis; ss.  
KW iron uptake; bioremediation; yellow stripe 1; ysl; gene; ss.

OS Arabidopsis sp.

PN WO200240688-A2.

PD 23-MAY-2002.

PF 16-NOV-2001; 2001WO-US43101.

PR 16-NOV-2000; 2000US-249222P.

PA (UYYA ) UNIV YALE.

PI Walker EL, Dellaporta S;

DR WPI; 2002-490144/52.

XX

PT New yellow stripe1 and yellow stripe1-like genes, useful for altering  
PT the distribution of iron within the plant body so that edible parts of  
PT crop plants have more iron, or for producing plants useful in enhancing  
PT iron uptake from soil -

PS Claim 1; Page 65-111; 187pp; English.

CC The invention relates to an isolated nucleic acid molecule (I), maize  
CC yellow stripe 1 (y1s1) or yellow strip1-like (y1s1) from Arabidopsis  
CC (ABN05763-ABN5771). (I) is useful for generating transgenic plants which  
CC can be used for enhancing iron uptake from soil and for bioremediation of  
CC metal or heavy metal contaminated soil. (I) may also be used to alter the  
CC distribution of iron within the plant body so that edible parts of crop  
CC plants have more iron. Transgenic plants may also be used in conventional  
CC plant breeding schemes to produce progeny which also contain the gene of  
CC interest. The present sequence is that of the Arabidopsis y1s1 encoding  
CC cDNA of the invention.

Sequence 82952 BP; 26971 A; 14261 C; 14365 G; 27355 T; 0 other;

Query Match	7.7%;	Score 48.2;	DB 24;	Length 82952;
-------------	-------	-------------	--------	---------------

Matches 134; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

281 TATTACTTTTATGCTGTACCGCATGGATTTCATTACGCTACACAAGTTGGTCATTAT 340

Db 21657 TACTACGTCGTGTTGGATCATATTGGGTTTCTATTGTTATTAAGCTACTTGAACATT 21598

341 GGACTACAAAGGTTTTTAATTAAATGGAGTAAATTTAATATATAATAAAAGGTTGAAAG 400

Db 21597 CGACCAACATCTCATTTAAAAACCTTTTCAATAGATAACAAACTAGAAAGAAG 21538

401 GAATGTTGCTTCTCAACAACACTTGTTCGACTTTTACGTAAGTTTACGTTCAAGTAATTG 460

Db 21537 CATCATATGCTTATCATCAAACCTTTAATCATTAAATTTCGGTAATTAAAGAATCT 21478

461 TTATAGGTATTGTTCCATATATGATTATAAAAGGAGAGTTTCAAACTTTTACCATTAAT 520

Db 21477 TTAACCTTCTGCTTATCAGAAATTTTTTTAAATGAAGAATCACTTACATATGTCCTTTAT 21418

521 ATTATGCAAGTCCCTTCAGGTGTGGCAGATTATATGCT 557

Db 21417 TAGATATATGCTTTTAAATATGCTTTCTCAATCCT 21381

```
Search completed: October '4, 2003, 04:04:50
Job time : 223 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 4, 2003, 03:48:37 ; Search time 56 Seconds  
(without alignments)  
4902.506 Million cell updates/sec

Title: US-10-010-160-1  
Perfect score: 622  
Sequence: 1 atgctgatgctccagctaa.....aattatcgctgcagatc 622

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents\_NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	46.6	7.5	8920	US-08-446-855A-1	Sequence 1, Appli
C 2	46.6	7.5	8920	US-09-150-741-1	Sequence 1, Appli
C 3	44.4	7.1	5181	US-08-257-073-10	Sequence 10, Appli
4	44.4	7.1	5852	US-07-867-106-2	Sequence 2, Appli
5	43.8	7.0	70000	US-09-851-896-3	Sequence 3, Appli
6	43.6	7.0	19124	US-08-487-826B-13	Sequence 13, Appli
7	42.8	6.9	66804	US-09-740-041-3	Sequence 3, Appli
C 8	41	6.6	3255	US-09-601-198-108	Sequence 108, App
C 9	40	6.4	19124	US-08-487-826B-13	Sequence 13, Appli
10	39.8	6.4	1400	US-08-305-764C-57	Sequence 57, Appli
11	39.8	6.4	1400	US-08-305-764C-55	Sequence 55, Appli
C 12	39.8	6.4	2095	US-08-305-764C-55	Sequence 55, Appli
C 13	39.8	6.4	9636	US-08-323-170B-1	Sequence 1, Appli
C 14	39.8	6.4	9636	US-08-954-441-1	Sequence 1, Appli
15	39	6.3	4526	US-07-855-412B-4	Sequence 4, Appli
16	39	6.3	4526	US-08-308-887A-4	Sequence 4, Appli
17	39	6.3	4526	US-08-881-094-4	Sequence 4, Appli
18	38.8	6.2	2861	US-08-299-953-1	Sequence 1, Appli
19	38.8	6.2	2861	US-08-459-415-1	Sequence 1, Appli
20	38.8	6.2	2861	US-09-066-687-1	Sequence 1, Appli
21	38.8	6.2	2861	PCT-US95-11231-1	Sequence 1, Appli
22	38.8	6.2	3881	US-08-299-953-2	Sequence 2, Appli
23	38.8	6.2	3881	US-08-459-415-2	Sequence 2, Appli
24	38.8	6.2	3881	US-09-066-687-2	Sequence 2, Appli
25	38.8	6.2	3881	PCT-US95-11231-2	Sequence 2, Appli
26	38.6	6.2	1664976	US-08-916-421B-1	Sequence 1, Appli
27	37.8	6.1	701	US-08-998-416-701	Sequence 701, App

28	37.8	6.1	1401	US-09-134-001C-595	Sequence 595, App
C 29	37.8	6.1	2142	US-09-107-532A-905	Sequence 905, App
C 30	37.8	6.1	2434	US-09-489-847-67	Sequence 67, Appl
31	37.6	6.0	1507	US-07-641-143B-1	Sequence 1, Appli
32	37.6	6.0	1507	US-08-124-290-1	Sequence 1, Appli
33	37.6	6.0	1507	US-08-696-372A-1	Sequence 1, Appli
34	37.4	6.0	1296	US-09-107-532A-3402	Sequence 3402, Ap
35	37.4	6.0	1875	US-08-286-325A-1	Sequence 1, Appli
36	37.4	6.0	3234	US-08-286-325A-7	Sequence 2508, Ap
37	37.2	6.0	576	US-09-134-001C-2508	Sequence 3, Appli
C 38	37.2	6.0	6124	US-08-213-419B-3	Sequence 7, Appli
39	37	5.9	1918	US-09-599-360B-7	Sequence 60, Appli
C 40	37	5.9	1918	US-09-599-360B-60	Sequence 60, Appli
C 41	37	5.9	8501	US-09-298-367B-6	Sequence 595, App
42	36.8	5.9	658	US-08-998-416-595	Sequence 1, Appli
C 43	36.8	5.9	6152	US-08-973-462-1	Sequence 1, Appli
44	36.6	5.9	767	US-08-998-416-472	Sequence 472, App
45	36.4	5.9	3618	US-07-872-678A-36	Sequence 36, Appl

ALIGNMENTS

RESULT 1  
US-08-446-855A-1/c  
Sequence 1, Application US/08446855A  
Patent No. 5849573  
GENERAL INFORMATION:  
APPLICANT: Stewart, Thomas S  
APPLICANT: Flores, Maria V  
APPLICANT: O'Sullivan, William J  
TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl  
NUMBER OF INVENTION: phosphate synthetase II  
CORRESPONDENCE ADDRESS:  
ADDRESS: Nixon & Vanderhye PC  
STREET: 1100 No. 5849573th Glebe Road, 8th Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446, 855A  
FILING DATE: 06-Jul-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mitchard, Leonard C  
REGISTRATION NUMBER: 29,009  
REFERENCE/DOCKET NUMBER: 47-80  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8920 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: genomic  
US-08-446-855A-1  
Query Match  
Best Local Similarity 7.5%; Score 46.6; DB 2; Length 8920;  
Matches 211; Conservative 0; Mismatches 214; Indels 6; Gaps 3;  
C 173 TTTTCTACTATATTTTACAGATCATTCGTTGAGCTTACAGACAGTCAGATATG 232  
DB 8835 TTTGCTGATTTTATTTTAAAAAATCATTAACATTATTAATTTATTAATTAAGAAAT 8776

Qy	233	CTTATATTATTTATTTGCTCAAGAGATGACATTTTATTTGATGCGCAATATTACTTTTA	292
Db	8775	TTATRCATTTTAAATGTTATATATTTTATTTTATTTTATTAATTAATTTATTTAAAT	8712
Qy	293	TTGCTGTTCAGCGCATGGAATTCATTACGTGTCACAAAGTTGGTCATTATGCACTACAAA-G	351
Db	8715	TATATAAATTTATTTGTTAAATGAAAAACCATTTTGTTTATCATATAGTAGTAATTAATTA	8655
Qy	352	GTTTTTAAATTTAAATGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	411
Db	8655	AAATTTGTAATATATACAAAAATTTATTAATAAAAAATTAATTAATTAATTAATTAATTAATTA	8599
Qy	412	TCTCACAAACAACCTGTTCGACTTTTAA---CGTAGTTTAGTTCAGAGTAATTTTATAGGT	468
Db	8595	TTTATATATAAACAATTTATTAATTTTAAACGTAATATATATTAATTAATTAATTAATTAATTAAT	8533
Qy	469	ATTGTTCCATATGATTTATTAATAAGAGAGTTTCAAACTTTTATCCATTATATATATGCA	528
Db	8535	AT	8476
Qy	529	AGTCTTCAAGTGTGGCAG--ATTATATGCTTAATACAGGATAGTACTTGTTTATATATA	586
Db	8475	TGTTGTTATGCAATMAAGTTTATATATAATCAATMAATAGTAATGTTAAATCGAATTAAT	8413
Qy	587	CGCTAATTCCT	597
Db	8415	GAATTAATTTCT	8405

## RESULT 2

```

US-09-150-741-1/c
Sequence 1, Application US/09150741
Patent No. 6183996
GENERAL INFORMATION:
APPLICANT: Stewart et al.
TITLE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl Phosphates
Patent No. 6183996
TITLE OF INVENTION: Synthetase II
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/150,741
EARLIER FILING DATE: 1998-09-10
EARLIER APPLICATION NUMBER: PL6380
EARLIER FILING DATE: 1992-12-16
EARLIER APPLICATION NUMBER: AU93/00617
EARLIER FILING DATE: 1993-12-02
EARLIER APPLICATION NUMBER: 08/446,855
EARLIER FILING DATE: 1995-07-06
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 8920
TYPE: DNA
ORGANISM: Plasmodium falciparum
US-09-150-741-1

```

Query Match : 7.5%; Score 46.6; DB 3; Length 8920;

Best Local Similarity 49.0%; Pred. No. 0.011;  
Matches 211; Conservative 0; Mismatches 214; Indels 6; Gaps 3

173 TTTTCTACTAATTTTACAGAATCAATTCGGTTTGAGGTTACAGCAGACGTCAATATG 232

Db 8835 TTTGGTGAATTTTAAATAATCAATTAACATTATATAATTATTAAGAAT 877

233 CTTTATTATTATGTGCTCAGAGATAGCTATTTTATTGATGCCAATTACTTTT 292

Db 8775 TTTACATTTAATGTTATATATTTTATTTTAAATAATTATTAAT 871

293 TTGCTTACGGCATGGATTTTCATTACGTGACAAATTGGTCATTATGCACTACAA-G 351

Db 8715 TATAAAATTATTGGTAAATGAAAAAACCAITTTGGTTATACATATGAGTGAAATTAATA 865

352 GTTTTAAATTAAATGAGTAAATTAAATAATAAAGGGTTGAAAGGAATGTTGCT 411

	Sequence	Position
Db	AAATTTGATATATATAACAAATAATTTATTAATAAATAATCAATATTAATTAATGATCATTA	8596
Oy	412 TCTCAACAAACACCTTGTGCACCTTAA---CGTAGTTAGTCAAGTAATGTATAGST	468
Db	8595 TTTAATATAAAACATTATATATATATTAATTAACGTAAATATTAATTAATTAATATATAT	8536
Oy	469 ATTGTTCATATATGATTAATAAAGAGAGTTTCAACCTTTTACCAATTAATTAATGCA	528
Db	8535 ATACCTTTTCA	8476
Oy	529 AGTCTTCAGGCTGTGCAG--ATATATGCTTAATAACAGAAATGTACTTGTTTATATA	586
Db	8475 TGTGTATTGGAATGAAAGAAAGTTAATAATAATACATAATATAGTATATTAATGCAATTAATT	8416
Oy	587 CGCTAATTCCT	597
Db	8415 GAATTAATTTCT	8405

### RESULT 3

US-08-257-073-10/c  
; Sequence 10, Application US/08257073  
; Patent No. 5766597

```

; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
;

```

APPLICANT: de Taisne, Charles  
APPLICANT: Tine, John A.  
STATE OF TENNESSEE: WILLIAM D.

1 TITLE OF INVENTION: MALMIRIA RECOMBINANT FOIAVIRUS VACCINE  
1  
1  
1 NUMBER OF SEQUENCES: 143  
1  
1 CORRESPONDENCE ADDRESS:

ADDRESSEE: Curtis, Morris & Safford, P.C  
STREET: 530 Fifth Avenue, 25th Floor

STATE: New York  
CITY: New York

COUNTRY: UNITED STATES OF AMERICA  
ZIP: 10036

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY.
;

```

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/257,073  
FILING DATE: 09-JUN-1994

CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 70 00 007 503

APPLICATION NUMBER: US 08/075,783  
FILING DATE: 11-JUN-1993  
PRIORITY APPLICATION DATA:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/852,305  
FILING DATE: 18-MAR-1992

FILING DATE: 10-MAR-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: IIS 07/672 183

ATTORNEY/AGENT INFORMATION:  
FILING DATE: 20-MAR-1991  
AFFIDAVIT NUMBER: 00 00000000

NAME: Frommer, William S.  
REGISTRATION NUMBER: 25.506

REFERENCE/DOCKET NUMBER: 454310-2570  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712

TELEX: 425066 CURTMS  
; INFORMATION FOR SEQ ID NO: 10

```

; SEQUENCE CHARACTERISTICS:
;     LENGTH: 5181 base pair

```

```

; type: nucleic acid
; strandedness: single

```

US-08-257-073-10

Query Match	7.1%;	Score 44;	DB 1;	Length 5181;
Best Local Similarity	53.5%;	Pred. No. 0.042;		

Matches 92; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 404 TGTTCCTTCAACAAACCTGTCGACTTACGAGTTCAGTAAATGTTA 463  
Db 2733 TTTTTCCTCTCATATATTTTTCGATGATTTTCCCTTTTGTATTAATA 2674

QY 464 TAGTATTTTCATATATATTAATAAGAGAGTTTCAACTTTTACCAATTAAT 523  
Db 2673 AATCATTTCTTTGATATATATCAAAAAAGATGTTGAATCATGTTCAATATATC 2614

QY 524 ATGAAGCTTCAGGCTGCGACATTAATGCTTAATACGAAATGACT 575  
Db 2613 ATATATGAATACATAGCAGATGTTATTTTGAATATTAATATATCT 2562

RESULT 4  
US-07-867-106-2  
Sequence 2, Application US/07867106  
Patent No. 5389526  
GENERAL INFORMATION:  
APPLICANT: Slade, Martin B  
APPLICANT: Chang, Andy C M  
APPLICANT: Williams, Keith L  
TITLE OF INVENTION: Improved Plasmid Vectors for Cellular  
TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESSES:  
ADDRESSES: Woodcock Washburn Kurtz Mackiewicz & No. 5389526  
STREET: One Liberty Place 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/867,106  
FILING DATE: 19920625  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PJ 7187  
APPLICATION NUMBER: PCT/AU90/00530  
FILING DATE: 02-NOV-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Feeney, Joanne Longo  
REGISTRATION NUMBER: 35,134  
REFERENCE/DOCKET NUMBER: RICE-0002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5852 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2378..5038  
NAME/KEY: CDS  
LOCATION: 2378..5038  
US-07-867-106-2

Query Match 7.1%; Score 44; DB 1; Length 5852;  
Best Local Similarity 47.2%; Pred. No. 0.043;  
Matches 134; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 311 TTTCATTACGTATCAAGTTGGTCATATGACTACAAAGTTTAAATTAATGA 370

Db 2088 TTTTAAATAAATTTAAATTTATCTCATATTAATTAATTAATTAATGAT 2147  
QY 371 GTAAATTAATTAATAAAGGTTGAAGAAAGTTTGCCTCCACAAACCTGTGC 430  
Db 2148 AATATATCAAAATATTAATGATTTTGGCATGACAAATTAATTAATTTTGAAT 2207  
QY 431 GACTTTACGATGTTTGAATCAAGTAATGTTATAGGTAATGTTCCATATTAATAA 490  
Db 2208 AATTTTATTTTATTTTATTTTAAATTTCTTTTATTTTATTTTAAATTTTAA 2267  
QY 491 AAGGAGTTTTCAACTTTTACCATTAATTAATGAAGTCTTCAGGTGCGAGATT 550  
Db 2268 AATTTTATTTTCCACACTTTCATTTATTTTATTTTATTTATTAATTAATTTAT 2327  
QY 551 ATATGCTTAATACGAAATAGTACTGTTTATTAATGACTAAT 594  
Db 2328 TTAATTTAAATTAATTAATGTTTGGTTAAATTTTAAATGAAGATT 2371

RESULT 5  
US-09-851-896-3  
Sequence 3, Application US/09851896  
Patent No. 6410325  
GENERAL INFORMATION:  
APPLICANT: C. Frank Bennett  
APPLICANT: Susan M. Freier  
APPLICANT: Andrew T. Watt  
TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI (CA2+-INDEPENDENT)  
FILE REFERENCE: RTS-0220  
CURRENT APPLICATION NUMBER: US/09/851,896  
CURRENT FILING DATE: 2001-05-08  
NUMBER OF SEQ ID NOS: 89  
SEQ ID NO 3  
LENGTH: 70000  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
US-09-851-896-3

Query Match 7.0%; Score 43.8; DB 4; Length 70000;  
Best Local Similarity 46.2%; Pred. No. 0.098;  
Matches 199; Conservative 0; Mismatches 217; Indels 15; Gaps 1;

QY 183 TATTTTACAGATCATTTGGCTTGAGGTTACGACGACGATGATATGCTTAATAT 242  
Db 23179 TAGTTTGTGTTAAATGTTGGTAAATTAAGCAGTGAAGCACTGGGTCACAGCTTTT 23238  
QY 243 TTATGTTGCTCAAGATAGCTATTTATGATGCCAATTAATCTTTTATGCTGTAC 302  
Db 23239 CTTCCTGAGAGACTTTTCATTTACAGCTTGATCTCATTTCTTTATTTGCTATTTAG 23298  
QY 303 GGCATGATTTTCATTAAGTATGATGATGATGATGATGATGATGATGATGATGAT 362  
Db 23299 GTTTTGATTTTCTCAAGGTTTCAATCTTGATGATGATGATGATGATGATGATGAT 23358  
QY 363 TAAATGAGTAAATTAATTAATTAATAAAGGTTGAAGAAAGATGTTGCTTCAACAAC 422  
Db 23359 TTCTTCTAGGCTTTCAATTAATTTGCA-----TATTAATGTTCAACAGTA 23403  
QY 423 ACTTGTGCACTTTAGTATGTTAGTTCAAGTAATGTTATGTTATGTTTCAATATAT 482  
Db 23404 ACCCTTAATATCTTTGAGTTCTGATATGTTGTTGAATATCTTTTGTATTTCT 23463  
QY 483 GATTATTAAGAGAGATTTTCAACTTTTACCAATTAATTAATTAATTAATTAATTAAT 542  
Db 23464 GATTTTGTGTTGGGTTTCTCTTTTGTGTTAGTATTAAGATTAATTAATTAATTT 23523  
QY 543 GGCAGATTAATGCTTAATACAGAAATAGTATGTTTATTAATAGCTAATTTCTATAGAC 602  
Db 23524 TGTATATCTTTTCAAAAACTTTATTTTGTGATCTTTTGTATGTTTGTTCATATTTTC 23583





SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 108  
LENGTH: 3255  
TYPE: DNA  
ORGANISM: Ureaplasma urealyticum  
US-09-601-198-108

Query Match 6.6%; Score 41; DB 4; Length 3255;  
Best Local Similarity 47.5%; Pred. No. 0.21;  
Matches 122; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 321 TGTACAACTGGTGAATATGACCTCAAGGTTTAAATTAATGAGTAATTTAA 380  
DB 1645 TTTACATATAAACTACATTAAGATATCCAGATTGATCTTAAAGATGA 1586  
QY 381 TATAATAAAGGTTGAAGGATGTTGCTTCAACAAACACTGTGCACTTTAG 440  
DB 1585 TTATTTCAATTAACAAATAATTAATGATTAATAATCTGTAAATGA 1526  
QY 441 TAGTTAGTCAAGTAATTTGATATGATTTGTCATATATGATTAAGAGAG 500  
DB 1525 AAATTTATTTTAAATATTTAGTTGTTAGATCAATCTCATCAAGAACTA 1466  
QY 501 TTCAACTTTTACCATTAATATGCAAGTCTTCAGGTGCGAGATTATANG 560  
DB 1465 GTGAACAAACCTTAATGTTAAGGTCTCAATTTTCAATGACAGCCCTT 1406  
QY 561 TACAGAAATAGTACTTG 577  
DB 1405 AAATTAATAAGTTATG 1389

## RESULT 9

US-08-487-826B-13/C  
Sequence 13, Application US/08487826B  
Patent No. 593827

GENERAL INFORMATION:  
APPLICANT: Sim, Kim L.  
APPLICANT: Chitnis, Chetan  
APPLICANT: Miller, Louis H.  
APPLICANT: Peterson, David S.  
APPLICANT: Su, Xin-zhan  
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobb Martens Olson & Bear  
STREET: 620 Newport Center Drive 16th Floor  
CITY: Newport Beach  
STATE: California  
COUNTRY: US

ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,826B  
FILING DATE: 10-SEP-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelien, Ned  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: NIH121.001CP1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19124 base pairs  
TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-487-826B-13

Query Match 6.4%; Score 40; DB 2; Length 19124;  
Best Local Similarity 45.1%; Pred. No. 0.63;  
Matches 148; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

QY 161 ATTTGAAACAATTTCTACTATATTTTACAGAAATCTTGGTTGAGTTACAG 220  
DB 15844 ATTTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15785  
QY 221 AGTCAGTATATCTTTATTTATTTATTTATTTATTTATTTATTTATTTAT 15725  
DB 15784 TTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15725  
QY 281 TATTACTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 340  
DB 15724 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 15665  
QY 341 GCACTCAAAAGTTTAAATTTAAATGAGTAATTTAATATATAAAGGTTGA 400  
DB 15664 TTTCTCTTTTGTATTTATTTTATTTTATTTATTTATTTTATTTATTT 15605  
QY 401 GAATGTTTGTCTTCACAAACACTGTTCACCTTTTACGATTTTACGATTTG 460  
DB 15604 AATTTTATTTTGAATATCTTTTCAATTTTATTTTATTTTATTTTATTT 15545  
QY 461 TTATAGTATTTGTCATATATGATTTAT 488  
DB 15544 TTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15517

## RESULT 10

US-08-305-764C-57  
Sequence 57, Application US/08305764C  
Patent No. 5856090

GENERAL INFORMATION:  
APPLICANT: Epstein, David M.  
TITLE OF INVENTION: DNA METHYLASE LINKING REACTION  
NUMBER OF SEQUENCES: 71  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE  
STREET: 10550 No. 5856090th Torrey Pines Road  
CITY: La Jolla  
STATE: California  
COUNTRY: USA

ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/305,764C  
FILING DATE: 09-SEP-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 440.0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 784-9397  
TELEFAX: (619) 784-9397  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1400 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1257  
US-08-305-764C-57

Query Match 6.4%; Score 39.8; DB 2; Length 1400;  
Best Local Similarity 48.5%; Pred. No. 0.34;  
Matches 110; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 308 GGATTCATTACGTCGACAAAGTTGGTCATTAAGCAATGAGTTTAAATTAAT 367  
DB 887 GCATTCGACGACATTTACAAAAGTTATCTTTAAAAAGATGATGTAACCTTCT 946  
QY 368 GGAGTAATTTAATTAATTAATAAGGTTGAAAGATGTTGCTTCAACAAACCTTG 427  
DB 947 TAATTGACAAAATACGACTGGGCGAGTTAAACATTAGTTCTACCTATCACAAATTC 1006  
QY 428 TTCGACTTTACGTAAGTTTGAATCAAGTAATGTTATAGGTAATGTTCCATATATGATTA 487  
DB 1007 AACGATTAACAGTAATCTTTGTTAAGATGAGAAACAGGTAACGCTTTTAAACACGA 1066  
QY 488 TAAAGGAGAGTTTCAAACTTTTACCATTAATTAATGCAAGTCT 534  
DB 1067 ATGAATGCAAGCTATTTATGGGTTTCCAAAAGATTTGTATTCCT 1113

RESULT 11  
US-08-305-764C-59  
Sequence 59, Application US/08305764C  
Patent No. 5856090  
GENERAL INFORMATION:  
APPLICANT: Epstein, David M.  
TITLE OF INVENTION: DNA METHYLASE LINKING REACTION  
NUMBER OF SEQUENCES: 71  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE  
STREET: 10550 No. 5856090th Torrey Pines Road  
CITY: La Jolla  
STATE: California  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/305,764C  
FILING DATE: 09-SEP-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 440.0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 784-2937  
TELEFAX: (619) 784-9399  
INFORMATION FOR SEQ ID NO: 59:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1400 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1257  
US-08-305-764C-59

Query Match 6.4%; Score 39.8; DB 2; Length 1400;  
Best Local Similarity 48.5%; Pred. No. 0.34;  
Matches 110; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 308 GGATTCATTACGTCGACAAAGTTGGTCATTAAGCAATGAGTTTAAATTAAT 367  
DB 887 GCATTCGACGACATTTACAAAAGTTATCTTTAAAAAGATGATGTAACCTTCT 946  
QY 368 GGAGTAATTTAATTAATTAATAAGGTTGAAAGATGTTGCTTCAACAAACCTTG 427  
DB 947 TAATTGACAAAATACGACTGGGCGAGTTAAACATTAGTTCTACCTATCACAAATTC 1006  
QY 428 TTCGACTTTACGTAAGTTTGAATCAAGTAATGTTATAGGTAATGTTCCATATATGATTA 487  
DB 1007 AACGATTAACAGTAATCTTTGTTAAGATGAGAAACAGGTAACGCTTTTAAACACGA 1066  
QY 488 TAAAGGAGAGTTTCAAACTTTTACCATTAATTAATGCAAGTCT 534  
DB 1067 ATGAATGCAAGCTATTTATGGGTTTCCAAAAGATTTGTATTCCT 1113

RESULT 12  
US-08-305-764C-55  
Sequence 55, Application US/08305764C  
Patent No. 5856090  
GENERAL INFORMATION:  
APPLICANT: Epstein, David M.  
TITLE OF INVENTION: DNA METHYLASE LINKING REACTION  
NUMBER OF SEQUENCES: 71  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE  
STREET: 10550 No. 5856090th Torrey Pines Road  
CITY: La Jolla  
STATE: California  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/305,764C  
FILING DATE: 09-SEP-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 440.0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 784-2937  
TELEFAX: (619) 784-9399  
INFORMATION FOR SEQ ID NO: 55:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2095 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 9..1952  
US-08-305-764C-55

Query Match 6.4%; Score 39.8; DB 2; Length 2095;  
Best Local Similarity 48.5%; Pred. No. 0.38;  
Matches 110; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 308 GGATTCATTACGTCGACAAAGTTGGTCATTAAGCAATGAGTTTAAATTAAT 367

Db 1582 GCATTTAGAGCATTTACAAAAAGTATCTTTAAAAAGATGATGTAACCTTCTT 1641  
QY 368 GAGTAATTTAATATATATAAAGGTTGAAGAGATGTTGCTTCAACAAACCTTG 427  
Db 1642 TAATGTACAAAANACAGACTGGGSCAGTTAAACATAGTTTCTACCTACACAAATTC 1701  
QY 428 TTGACCTTTACGTAAGTTAGTTCAAGTAAATGTTATAGTATGTTCCATATATGATTA 487  
Db 1702 AACGATTAACAGAGTACTTTGTTAAGATGAGAAACAGGATCCGCTTTTAAACAGCA 1761  
QY 488 TAAAGAGAGTTTCAACCTTTTACCATATATTTGCAAGTCTT 534  
Db 1762 ATGAATGCAAGCTATTATAGGTTTCCAAAAGATTTGTTATTCCT 1808

## RESULT 13

US-08-323-170B-1/C  
Sequence 1, Application US/08323170B

Patent No. 5733772

GENERAL INFORMATION:

APPLICANT: Williamson, Kim C.

APPLICANT: Kaslow, David C.

TITLE OF INVENTION: Cloning and Expression of Plasmidium

TITLE OF INVENTION: Falliciparum Transmission-Blocking Target Antigen, Pfs230

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/323,170B

FILING DATE: 13-OCT-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/010,409

FILING DATE: 29-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Quine, Jonathan A.

REGISTRATION NUMBER: P-41,261

REFERENCE/DOCKET NUMBER: 015280-113100US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 9636 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 149..9556

US-08-323-170B-1

Query Match 6.4%; Score 39.8; DB 1; Length 9636;

Best Local Similarity 44.4%; Pred. No. 0.58; Indels 0; Gaps 0;

Matches 161; Conservative 0; Mismatches 202; Indels 0; Gaps 0;

QY 143 CAGCGTAATGAGAGTATTTGAACAATTTCTACTATATTTTACAGAAATCATTTG 202

Db 4973 CATCTGATCTGCGGTTTCTTTTACCTTAATCATGAATTAATTTTAC 4914

QY 203 GGTGAGGTTACGACAGCTCATATATGCTTTATTTATTTATGTTGCTCAAGATAG 262

Db 4913 GTTATGAGATTAATGTAATTCATATATGTCATTAAGATATTTATTTGATCAGTCCAG 4854  
QY 263 CTAATTTATGATGCCAATATATCTTTTATGCTGTAGCGATGATTTCAATGCTG 322  
Db 4853 TAACATATATATCAAAATTTATATTTTATTTTATTAAGACAGCAAGAACTTAGTGC 4794  
QY 323 TACAGTGTGATTAATGATGACAAAGGTTTAAATTTAAATGAGATTAATTTATA 382  
Db 4793 CTAAATGATTTGACATTTATATTTTATTTTATTTTAAATCAACATATATCAATTAATG 4734  
QY 383 TAATTAAGGTTGAAGAGATGTTGCTTCTCAACAAACACTGTTGACTTTACGTA 442  
Db 4733 ATATATCATTTATCCATTAAGCTGTATTTCCATATATTAATTAATGATTTATATGTA 4674  
QY 443 GTTATGATCAATATTTATATGATGATGTTGTCATATATATATTAATTAAGAGATTTT 502  
Db 4673 AGTTAGCCAAATTTATTTTAAATTAAGTATACCATTAATATTTGTTAAATTTGTATGTA 4614  
QY 503 CAA 505  
Db 4613 CTA 4611

## RESULT 14

US-08-954-441-1/C  
Sequence 1, Application US/08954441

Patent No. 6316000

GENERAL INFORMATION:

APPLICANT: Williamson, Kim C.

APPLICANT: Kaslow, David C.

TITLE OF INVENTION: Cloning and Expression of Plasmidium

TITLE OF INVENTION: Falliciparum Transmission-Blocking Target Antigen, Pfs230

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/954,441

FILING DATE: 20-OCT-1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/323,170

FILING DATE: 13-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/010,409

FILING DATE: 29-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Einhorn, Gregory P.

REGISTRATION NUMBER: 38,440

REFERENCE/DOCKET NUMBER: 015280-113110US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 9636 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 149..9556

US-08-954-441-1

Query Match 6.4%; Score 39.8; DB 4; Length 9636;  
Best Local Similarity 44.4%; Pred. No. 0.58;  
Matches 161; Conservative 0; Mismatches 202; Indels 0; Gaps 0;

QY 143 CAGGGTAAAGGAGTCACTTTTGAACAACTTTCTACTATTTTTCACAGAACTATTTC 202  
DB 4973 CATCTGATCTGCTGGTTTTTTTACCTAAATCAAGAAATATCTTTAAATATAC 4914  
QY 203 GGTTCAGGTACACACAGTCAGTATATGCTTTATTTATTTATGCTCAAGAGATAG 262  
DB 4913 GTTATGAGATAAATGTAATCATATATGTCATAGAAATTTATTTTATGATGAGCCAG 4854  
QY 263 CATTTTATGATGCAATATTTACTTTTATGCTGTTACGCGATGATTTTATTACGTG 322  
DB 4853 TAACTATATTAATCAAAATTTTATTTATTTATTTAAAGACAGTAAACCTTTAGGTC 4794  
QY 323 TACAGTTGTCATTTATGACCTACAAAGCTTTTAAATTTAATGAGTAAATTTATA 382  
DB 4793 CTAAATGATTTGACATTTATTTTCTTTTAAATCAACATATACATTTAAATG 4734  
QY 383 TAAATAAAGGTTGAAGAAATGTTGCTTCAACAAACACTTGTTCACATTTTACGTA 442  
DB 4733 ATATTCATTTATCCATTAAGCTGATTTCCATATTAATTAATGATTTATATGTA 4674  
QY 443 GTTTCAGTCAATATTTGTTATAGTATTTCCATATATGATTTAAAGAGAGCTTT 502  
DB 4673 AGTTAGCCAAATTTATTTTAAATTAAGTATCCATTAATTTGTTAAATTTGATGTA 4614  
QY 503 CAA 505  
DB 4613 CTA 4611

## RESULT 15

US-07-855-412B-4  
; Sequence 4, Application US/07855412B  
; Patent No. 5378819

## GENERAL INFORMATION:

APPLICANT: Ryan, C.A.; McGurl, B.F.; Pearce, G.L.

TITLE OF INVENTION: "SYSTEMIN"

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Harness, Dickey & Pierce

STREET: P.O. Box 828

CITY: Troy

STATE: Michigan

COUNTRY: USA

ZIP: 48303

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette-5.25 inch, 1.2mb storage

COMPUTER: IBM PC/386 Compatible

OPERATING SYSTEM: MS-DOS 4.01

SOFTWARE: Word for Windows-t

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/855,412B

FILING DATE: 03/19/92

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Deann F.

REGISTRATION NUMBER: 36,683

REFERENCE/DOCKET NUMBER: 755500002

TELEPHONE: (810) 641-1600

TELEFAX: (810) 641-0270

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 4526 bases

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

DESCRIPTION: prosystemin genomic DNA, Figures 8A-8C.

US-07-855-412B-4  
Query Match 6.3%; Score 39; DB 1; Length 4526;  
Best Local Similarity 52.8%; Pred. No. 0.75;  
Matches 84; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 196 TCATTTCGTTTGAAGTTACAGACAGTCAGTATATGCTTTATTTATTTATGTTCTCAA 255  
DB 2148 TAAATTTATTTTATTTTTCAGATTTTATAGTAAAGGATTTTATTTTCAAAAA 2207  
QY 256 GAGATAGCTATTTATGATGCAATATTTACTTTTATGCTGTTACGCGATGATTTCA 315  
DB 2208 ATGGAATCATTTTCAAGAAATTAATTTTATTTTGGTAACCTTTAATCTGATATATAT 2267  
QY 316 TTACGTGACAGTTGTCATTTATGACCTACAAAGTT 354  
DB 2268 TCTCAGAGATGATGATGATGCAAGAAATAAGTT 2306

## RESULT 16

US-08-308-887A-4  
; Sequence 4, Application US/08308887A  
; Patent No. 5883076

## GENERAL INFORMATION:

APPLICANT: Ryan, C.A.

APPLICANT: McGurl, B.F.

APPLICANT: Pearce, G.L.

TITLE OF INVENTION: "SYSTEMIN"

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Harness, Dickey & Pierce, P.L.C.

STREET: P.O. Box 828

CITY: Troy

STATE: Michigan

COUNTRY: USA

ZIP: 48303

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette-5.25 inch, 1.2mb storage

COMPUTER: IBM PC/386 Compatible

OPERATING SYSTEM: MS-DOS 4.01

SOFTWARE: Word for Windows-t

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/308,887A

FILING DATE: September 19, 1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Smith, Deann F.

REGISTRATION NUMBER: 36,683

REFERENCE/DOCKET NUMBER: 755500001USA

TELEPHONE: (810) 641-1600

TELEFAX: (810) 641-0270

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 4526 bases

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

DESCRIPTION: prosystemin genomic DNA, Figures 8A-8C.

US-08-308-887A-4  
Query Match 6.3%; Score 39; DB 2; Length 4526;  
Best Local Similarity 52.8%; Pred. No. 0.75;  
Matches 84; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 196 TCATTTCGTTTGAAGTTACAGACAGTCAGTATATGCTTTATTTATTTATGTTCTCAA 255  
DB 2148 TAAATTTATTTTATTTTTCAGATTTTATAGTAAAGGATTTTATTTTCAAAAA 2207  
QY 256 GAGATAGCTATTTATGATGCAATATTTACTTTTATGCTGTTACGCGATGATTTCA 315

Db 2208 ATGCAATCATTTTCAGAGTAAATATTTTGGTAACCTTATCTGATATATAT 2267  
Qy 316 TTACGTGACAGTGTGGCATTATGACCTACAAGGTT 354  
Db 2268 TCTCCAGAGATGATGATGACACAGAAAACCTAAAGGTT 2306

## RESULT 17

US-08-094-4  
Sequence 4, Application US/08881094A  
Patent No. 6022739

## GENERAL INFORMATION:

APPLICANT: Ryan, Clarence A  
APPLICANT: Pearce, Gregory L  
APPLICANT: McGurk, Barry F  
TITLE OF INVENTION: Systemin  
FILE REFERENCE: 7555-00001CPB  
CURRENT FILING DATE: 1997-07-09  
EARLIER FILING DATE: 1994-09-19  
EARLIER FILING DATE: 1993-03-18  
EARLIER FILING DATE: 1992-03-19  
EARLIER FILING DATE: 1990-05-25  
EARLIER FILING DATE: 1991-05-24  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 4  
LENGTH: 4526  
TYPE: DNA  
ORGANISM: Lycopersicon esculentum  
US-08-881-094-4

Query Match 6.3%; Score 39; DB 3; Length 4526;  
Best Local Similarity 52.8%; Pred. No. 0.75; Indels 0; Gaps 0;

Matches 84; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy 196 TCATTTCGTTGAGTTACAGCAGTCAATATGCTTATTTATTTATGCTCA 255  
Db 2148 TAAATTTATTTTATTTTTCAGATTTTATAGAGGATTTTATTTTTCAAAA 2207  
Qy 256 GAGATGCTATTTTATGATGCAATATTTACTTTTATGCTGTTACGCGATGTTCA 315  
Db 2208 ATGCAATCATTTTCAGAGTAAATATTTTGGTAACCTTATCTGATATATAT 2267  
Qy 316 TTACGTGACAGTGTGGCATTATGACCTACAAGGTT 354  
Db 2268 TCTCCAGAGATGATGATGACACAGAAAACCTAAAGGTT 2306

## RESULT 18

US-08-239-953-1  
Sequence 1, Application US/08299953  
Patent No. 5646333

## GENERAL INFORMATION:

APPLICANT: Dobres, Michael S. and Mandaci, Sevnuur  
TITLE OF INVENTION: A Plant Promoter Useful for Directing the  
TITLE OF INVENTION: Expression of Foreign Proteins to the Plant Epidermis  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: Woodcock Washburn Kurtz Mackiewicz & No. 5646333xis  
STREET: One Liberty Place 46th. Floor  
CITY: Philadelphia  
STATE: PA  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/299,953  
FILING DATE: Herewith  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Beardell, Lori Y.  
REGISTRATION NUMBER: 34,293  
REFERENCE/DOCKET NUMBER: NOVA-0003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-564-8960  
TELEFAX: 215-568-3439

## INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 2861 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-239-953-1

Query Match 6.2%; Score 38.8; DB 1; Length 2861;  
Best Local Similarity 44.7%; Pred. No. 0.74; Indels 0; Gaps 0;  
Matches 151; Conservative 0; Mismatches 187; Indels 0; Gaps 0;

Qy 175 TTCTACTATATTTTACAGATCAATTCGTTGAGTTACAGCAGTCAATATGCT 234  
Db 1108 TTTTACATGTTAAAGAAATTAATATGATTTTATTAATCACTGAAATTTATAT 1167  
Qy 235 TTATTTATTTATGTTCCAGAGATGCTATTTATGATGCAATTTATCTTTAT 294  
Db 1168 TATCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1227  
Qy 295 GCTGTACGCGATGATTTTCACTTACGTACAGTGTGGCATTATGACCTACAAGGTT 354  
Db 1228 TTTTGAATATAATTTTATTAATGATGTAATAAATCTAAATTAATTAATTAAT 1287  
Qy 355 TTTTAATTTAAATGAGTAAATTTAATTAATTAATTAATTAATTAATTAATTAAT 414  
Db 1288 TTTTAAAGTCATTTTATTTTGGATTTTGGACATATCAAGGAAGCTAATCATTAAT 1347  
Qy 415 CAACAACACTGTTTCAGCTTTTACGTATGATGCAAGTATGATAGTATGTT 474  
Db 1348 TTTCAAAATTAATTTCTTTTATTAATTAATTAATTAATTAATTAATTAATTAAT 1407  
Qy 475 CCATATATGATTAATAAGAGAGGTTTCAACTTTT 512  
Db 1408 CTTTTCCTCTCTCCATCATTTATCTAGATTAATTT 1445

## RESULT 19

US-08-459-415-1  
Sequence 1, Application US/08459415  
Patent No. 5744334

## GENERAL INFORMATION:

APPLICANT: Dobres, Michael S. and Mandaci, Sevnuur  
TITLE OF INVENTION: A Plant Promoter Useful for Directing the  
TITLE OF INVENTION: Expression of Foreign Proteins to the Plant Epidermis  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: Woodcock Washburn Kurtz Mackiewicz & No. 5744334xis  
STREET: One Liberty Place 46th. Floor  
CITY: Philadelphia  
STATE: PA  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,415  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/299,953  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Beardell, Lori Y.  
REGISTRATION NUMBER: 34,293  
REFERENCE/DOCKET NUMBER: NOVA-0003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-564-8960  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2861 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-459-415-1

Query Match  
Best Local Similarity 44.7%; Score 38.8; DB 1; Length 2861;  
Pred. No. 0.74; Mismatches 187; Indels 0; Gaps 0;  
Matches 151; Conservative 0;

175 TTCTACTATATTTTACAGATCATTTGGTTGAGGTACAGACAGTCAGTATATGCT 234  
1108 TTTTACAAATGTTAAAGAAATTAATTAATTTTATTAATTAATTAATTAATTAAT 1167  
235 TTAATTTATTAATGTTGCTCAAGATAGCTATTTATGATGCAATTAATTTTATT 294  
1168 TATCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1227  
295 GCTGTACGCGATGATTTTCACTGATGATGATGATGATGATGATGATGATGATGAT 354  
1228 TTTTGAATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1287  
355 TTTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 414  
1288 TTTTAAAGTCAATTTTGTGATTTTGTGACATATCAAGAGATATATATTTATTT 1347  
415 CAACAACACTGTTGACCTTTTACGATTTTACGATTTTACGATTTTACGATTTTACG 474  
1348 TTTCAACAATTAATTTCTTTTACTTAATTAATTAATTAATTAATTAATTAATTAAT 1407  
475 CCATATATGATTTAATTAAGAGAGTTTCAACTTTT 512  
1408 CTTTTCCTCTCTCCCAATCATTTATCTAGAGTAAATTT 1445

RESULT 20  
US-09-066-687-1  
Sequence 1, Application US/09066687  
Patent No. 6339185  
GENERAL INFORMATION:  
APPLICANT: Dobres, Michael S. and Mandaci, Sevnuur  
TITLE OF INVENTION: A Plant Promoter Useful for Directing the  
TITLE OF INVENTION: Expression of Foreign Proteins to the Plant Epidermis  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6339185ctis  
STREET: One Liberty Place 46th. Floor  
CITY: Philadelphia  
STATE: PA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/066,687  
FILING DATE: Herewith  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Beardell, Lori Y.  
REGISTRATION NUMBER: 34,293  
REFERENCE/DOCKET NUMBER: NOVA-0003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-564-8960  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2861 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-09-066-687-1

Query Match  
Best Local Similarity 44.7%; Score 38.8; DB 4; Length 2861;  
Pred. No. 0.74; Mismatches 187; Indels 0; Gaps 0;  
Matches 151; Conservative 0;

175 TTCTACTATATTTTACAGATCATTTGGTTGAGGTACAGACAGTCAGTATATGCT 234  
1108 TTTTACAAATGTTAAAGAAATTAATTAATTTTATTAATTAATTAATTAATTAAT 1167  
235 TTAATTTATTAATGTTGCTCAAGATAGCTATTTATGATGCAATTAATTTTATT 294  
1168 TATCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1227  
295 GCTGTACGCGATGATTTTCACTGATGATGATGATGATGATGATGATGATGATGAT 354  
1228 TTTTGAATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1287  
355 TTTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 414  
1288 TTTTAAAGTCAATTTTGTGATTTTGTGACATATCAAGAGATATATATTTATTT 1347  
415 CAACAACACTGTTGACCTTTTACGATTTTACGATTTTACGATTTTACGATTTTACG 474  
1348 TTTCAACAATTAATTTCTTTTACTTAATTAATTAATTAATTAATTAATTAATTAAT 1407  
475 CCATATATGATTTAATTAAGAGAGTTTCAACTTTT 512  
1408 CTTTTCCTCTCTCCCAATCATTTATCTAGAGTAAATTT 1445

RESULT 21  
PCT-US95-11231-1  
Sequence 1, Application PCT/US9511231  
GENERAL INFORMATION:  
APPLICANT: Dobres, Michael S. and Mandaci, Sevnuur  
TITLE OF INVENTION: A Plant Promoter Useful for Directing the Expression  
TITLE OF INVENTION: of Foreign Proteins to the Plant Epidermis  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris  
STREET: One Liberty Place 46th. Floor  
CITY: Philadelphia  
STATE: PA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/11231

FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/299,953  
FILING DATE: September 2, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Beardell, Lori Y.  
REGISTRATION NUMBER: 34,293  
REFERENCE/DOCKET NUMBER: NOVA-0016  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-564-8960  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2861 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
PCT-US95-11231-1

Query Match 6.2%; Score 38.8; DB 5; Length 2861;  
Best Local Similarity 44.7%; Pred. No. 0.74;  
Matches 151; Conservative 0; Mismatches 187; Indels 0; Gaps 0;

Qy 175 TTCTACTATATTTTACAGATCATTCGTTGAGTTACAGACAGTCAATATGCT 234  
Db 1108 TTTTACATGTTAAAGAAATTAATATAGTTTATTAATTAATCTAGAAATTTATAT 1167  
Qy 235 TTATTTATTTATGTTGTCAAGATGCTATTTATGATGCCAATATTAATCTTTTAT 294  
Db 1168 TATCTTTTATTAATATTAATTAATTAATCTTCCATTTTATTAATTAATTAAGATTT 1227  
Qy 295 GCTGTACGAGATGATTTCAATACGTGTACAGAGTTGCTGATTAAGCAAAAGTT 354  
Db 1228 TTTTGTATTAATTTTATTAATGCTGTAAGAAATCTAAATATTAATTAATCTCCGCTCT 1287  
Qy 355 TTTAATTTAATGAGTAATTAATTAATTAATTAAGGTTGAAGAAATGTTGCTCT 414  
Db 1288 TTTTAAGTCATTTTATTTGATTTTGAATATCAAGAAAGCTAATCAATATTTGTTAT 1347  
Qy 415 CACAAACACTGTTGCACTTTTACGTAGTTTCAAGTAATGTTATGTAATGTT 474  
Db 1348 TTTCAACAATATTTCTTTTACTTAATATCCCTTAATTAATTAATTAATCACTTTA 1407  
Qy 475 CCATATATGATTAATTAAGAGAGTTTCAAACTTTT 512  
Db 1408 CTTTTCCTCTCTCCATCAATCAATATCTAGATTAATTT 1445

RESULT 22  
US-08-299-953-2  
Sequence 2, Application US/08299953  
Patent No. 5646333

GENERAL INFORMATION:  
APPLICANT: Dobres, Michael S. and Mandaci, Sevnu  
TITLE OF INVENTION: A Plant Promoter Useful for Directing the  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: Woodcock Washburn Kurtz Mackiewicz & No. 5646333r1s  
STREET: One Liberty Place 46th. Floor  
CITY: Philadelphia  
STATE: PA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/299,953  
FILING DATE: Herewith  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Beardell, Lori Y.  
REGISTRATION NUMBER: 34,293  
REFERENCE/DOCKET NUMBER: NOVA-0003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-564-8960  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3881 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-299-953-2

Query Match 6.2%; Score 38.8; DB 1; Length 3881;  
Best Local Similarity 44.7%; Pred. No. 0.81;  
Matches 151; Conservative 0; Mismatches 187; Indels 0; Gaps 0;

Qy 175 TTCTACTATATTTTACAGATCATTCGTTGAGTTACAGACAGTCAATATGCT 234  
Db 1108 TTTTACATGTTAAAGAAATTAATATAGTTTATTAATTAATCTAGAAATTTATAT 1167  
Qy 235 TTATTTATTTATGTTGTCAAGATGCTAATTTATGATGCCAATATTAATCTTTTAT 294  
Db 1168 TATCTTTTATTAATATTAATTAATTAATCTTCCATTTTATTAATTAATTAAGATTT 1227  
Qy 295 GCTGTACGAGATGATTTCAATACGTGTACAGAGTTGCTGATTAAGCAAAAGTT 354  
Db 1228 TTTTGTATTAATTTTATTAATGCTGTAAGAAATCTAAATATTAATTAATCTCCGCTCT 1287  
Qy 355 TTTAATTTAATGAGTAATTAATTAATTAATTAAGGTTGAAGAAATGTTGCTCT 414  
Db 1288 TTTTAAGTCATTTTATTTGATTTTGAATATCAAGAAAGCTAATCAATATTTGTTAT 1347  
Qy 415 CACAAACACTGTTGCACTTTTACGTAGTTTCAAGTAATGTTATGTAATGTT 474  
Db 1348 TTTCAACAATATTTCTTTTACTTAATATCCCTTAATTAATTAATTAATCACTTTA 1407  
Qy 475 CCATATATGATTAATTAAGAGAGTTTCAAACTTTT 512  
Db 1408 CTTTTCCTCTCTCCATCAATCAATATCTAGATTAATTT 1445

RESULT 23  
US-08-459-415-2  
Sequence 2, Application US/08459415  
Patent No. 5744334

GENERAL INFORMATION:  
APPLICANT: Dobres, Michael S. and Mandaci, Sevnu  
TITLE OF INVENTION: A Plant Promoter Useful for Directing the  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: Woodcock Washburn Kurtz Mackiewicz & No. 5744334r1s  
STREET: One Liberty Place 46th. Floor  
CITY: Philadelphia  
STATE: PA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,415  
FILING DATE: 02-JUN-1995

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/299,953  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Beardell, Lori Y.  
REGISTRATION NUMBER: 34,293  
REFERENCE/DOCKET NUMBER: NOVA-0003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-564-8960  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3881 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-459-415-2

Query Match 6.2%; Score 38.8; DB 1; Length 3881;  
Best Local Similarity 44.7%; Pred. No. 0.81;  
Matches 151; Conservative 0; Mismatches 187; Indels 0; Gaps 0;

QY 175 TTCTACTATATTTTACAGATCATTCGGTTGAGGTTACGACAGTCAGTATATGCT 234  
DB 1108 TTTTACATGTTAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1167

QY 235 TTATTTATTTAGTGTCTCAAGAGATGCTATTTATTTAGTCAATTTACTTTTAT 294  
DB 1168 TATCTTTTATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1227

QY 295 GCTGTACGAGATGATTTTCAAGATGCTATTAAGTGTGATTTGATTTGATTTGAT 354  
DB 1228 TTTTGAATATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1287

QY 355 TTTAAATTTAAATGAGTAATTTATTAATTAATTAATTAATTAATTAATTAATTAAT 414  
DB 1288 TTTTAAAGTCAATTTTGTGATTTTGCACATATCAAGAGATCAATCAATTTATTT 1347

QY 415 CAACAACACTGTTGCACTTTAGTGTGATTTAGTCAAGTATTTAGTATTTGTT 474  
DB 1348 TTTTCAACAATTAATTTCTTTTACTTAATTAATTAATTAATTAATTAATTAATTAAT 1407

QY 475 CCATATATGATTTATTAAGAGAGTTTCAACTTTT 512  
DB 1408 CTTTCTCTCTCTCCATCAATCAATTAATTAATTAATTAATTAATTAATTAAT 1445

RESULT 24  
US-09-066-687-2  
Sequence 2, Application US/0906687  
Patent No. 6339185  
GENERAL INFORMATION:  
APPLICANT: Dobres, Michael S. and Mandaci, Sevnur  
TITLE OF INVENTION: A Plant Promoter Useful for Directing the  
TITLE OF INVENTION: Expression of Foreign Proteins to the Plant Epidermis  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6339185r1s  
STREET: One Liberty Place 46th. Floor  
CITY: Philadelphia  
STATE: PA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/066,687

FILING DATE: Herewith  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Beardell, Lori Y.  
REGISTRATION NUMBER: 34,293  
REFERENCE/DOCKET NUMBER: NOVA-0003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-564-8960  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3881 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-09-066-687-2

Query Match 6.2%; Score 38.8; DB 4; Length 3881;  
Best Local Similarity 44.7%; Pred. No. 0.81;  
Matches 151; Conservative 0; Mismatches 187; Indels 0; Gaps 0;

QY 175 TTCTACTATATTTTACAGATCATTCGGTTGAGGTTACGACAGTCAGTATATGCT 234  
DB 1108 TTTTACATGTTAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1167

QY 235 TTATTTATTTAGTGTCTCAAGAGATGCTATTTATTTAGTCAATTTACTTTTAT 294  
DB 1168 TATCTTTTATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1227

QY 295 GCTGTACGAGATGATTTTCAAGATGCTATTAAGTGTGATTTGATTTGATTTGAT 354  
DB 1228 TTTTGAATATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1287

QY 355 TTTAAATTTAAATGAGTAATTTATTAATTAATTAATTAATTAATTAATTAATTAAT 414  
DB 1288 TTTTAAAGTCAATTTTGTGATTTTGCACATATCAAGAGATCAATCAATTTATTT 1347

QY 415 CAACAACACTGTTGCACTTTTACGTAATTTAGTCAAGTATTTGATTAATTTGTT 474  
DB 1348 TTTTCAACAATTAATTTCTTTTACTTAATTAATTAATTAATTAATTAATTAATTAAT 1407

QY 475 CCATATATGATTTATTAAGAGAGTTTCAACTTTT 512  
DB 1408 CTTTCTCTCTCTCCATCAATCAATTAATTAATTAATTAATTAATTAATTAAT 1445

RESULT 25  
PCT-US95-11231-2  
Sequence 2, Application PC/TUS9511231  
GENERAL INFORMATION:  
APPLICANT: Dobres, Michael S. and Mandaci, Sevnur  
TITLE OF INVENTION: A Plant Promoter Useful for Directing the Expression  
TITLE OF INVENTION: of Foreign Proteins to the Plant Epidermis  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris  
STREET: One Liberty Place 46th. Floor  
CITY: Philadelphia  
STATE: PA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/11231  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:



APPLICATION NUMBER: 08/299,953  
FILING DATE: September 2, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Beardsell, Lori Y.  
REGISTRATION NUMBER: 34,293  
REFERENCE/DOCKET NUMBER: NOVA-0016  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3439  
TELEFAX: 215-568-8960  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3881 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
PCT-US95-11231-2

Query Match 6.2%; Score 38.8; DB 5; Length 3881;  
Best Local Similarity 44.7%; Pred. No. 0.81; Mismatches 187; Indels 0; Gaps 0;  
Matches 151; Conservative 0; Mismatches 187; Indels 0; Gaps 0;

QY 175 TTCTACTATATTTTACAGATTCGTTGAGTTGACAGACAGTATATGCT 234  
DB 1108 TTTTACATGTTAAAGATTAATATGTTTATATATGCTAGAAATTTATAT 1167  
QY 235 TTATTTATTTGTTGCTCAAGATAGCTATTTATGATGCCAATTTACTTTTAT 294  
DB 1168 TATCTTTTATTTTATTAATAAATTTATCTTCATTTTATTAATAATAAGATAT 1227  
QY 295 GCTGTACGCGATGATTTCTATGCTGACAGTTGGTGCATTATGACTACAAAGTT 354  
DB 1228 TTTTGTATATATTTTATTAAGCTGTAATAAATCTAAATATTAATCTCCCTGTTCT 1287  
QY 355 TTTAAATTTAAATGAGTAAATTTAATAATAATAAAGGTTGAAAGATTTGCTCT 414  
DB 1288 TTTTAAGTGTATTTTATGATTTTGGACATATCAAGAGATTAATCATTTATTTAT 1347  
QY 415 CAAGAAACACTGTTGCACTTTTACGTTAGTTAGTCAAGTAATTTGTTATGATTT 474  
DB 1348 TTTCAACCAATTAATTTCTTTTACTTATTAATACCTTAATTTATTTATTCATCACTTTA 1407  
QY 475 CCATATATGATTATAAAGAGAGTTTCAACTTTT 512  
DB 1408 CTTTCTCTCTCTCCATTCATTATCTAGAGATTAATTT 1445

RESULT 26  
US-08-916-421B-1  
Sequence 1, Application US/08916421B  
Patent No. 6503729  
GENERAL INFORMATION:  
APPLICANT: Bull et al.  
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus  
Patent No. 6503729  
TITLE OF INVENTION: jannaschii  
FILE REFERENCE: PB275  
CURRENT APPLICATION NUMBER: US/08/916,421B  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: US 60/024,428  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 1664976  
TYPE: DNA  
ORGANISM: Methanococcus jannaschii  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (28222)..(28222)  
OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature  
LOCATION: (28257)..(28258)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (84773)..(84773)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (84808)..(84808)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (84812)..(84812)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (98120)..(98120)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (98159)..(98159)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (98239)..(98239)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (98266)..(98266)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (103968)..(103968)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (148948)..(148948)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (163385)..(163385)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (191989)..(191989)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (191995)..(191995)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (231980)..(231980)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234187)..(234187)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234220)..(234220)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234814)..(234814)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (309398)..(309398)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (309418)..(309418)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (312837)..(312837)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (312993)..(312993)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (319226)..(319226)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (559167)..(559167)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature

LOCATION: (559241)..(559241)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (600992)..(600992)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (622708)..(622708)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (657081)..(657081)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (657203)..(657203)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (674435)..(674435)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (682442)..(682442)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (713652)..(713652)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (741684)..(741684)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (779455)..(779455)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (779676)..(779676)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (855539)..(855539)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (871619)..(871619)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1084830)..(1084830)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1096846)..(1096846)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1119881)..(1119881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1130881)..(1130881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1310988)..(1310988)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1313224)..(1313224)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1349473)..(1349473)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1349491)..(1349491)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1470091)..(1470091)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1569020)..(1569020)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1602912)..(1602912)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1603734)..(1603734)

OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1637998)..(1637998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1664854)..(1664854)  
OTHER INFORMATION: n equals a, t, c, or g  
US-08-916-421B-1  
Query Match 6.2%; Score 38.6; DB 4; Length 1664976;  
Best Local Similarity 48.4%; Pred. No. 5;  
Matches 107; Conservative 0; Mismatches 114; Indels 0; Gaps 0;  
Qy 323 TACAGTTGGTCACTATGAGCTTCAAGGTTTAAATTTAAATGAGTAATTTATA 382  
Db 1179668 TAGAGTTGATGAGGCAAGATTAAGACCTTTAAAGATGAGAACTAATACATG 1179727  
Qy 383 TAATMAAGGTTGAAGAAATGTTGCTTCTCAACAACACTTGTTCAGCTTTACGTA 442  
Db 1179728 ATGAAGAAGTCATTAAGGAATATTTTAGCAGCAGCTTAATTTTAAAGATGCTGTA 1179787  
Qy 443 GTTAGTTCAAGTAATGTTATAGTATGTCATATATGATTAATAAGAGATTTT 502  
Db 1179788 AAAAGCTTGAATTAATTTATTTGCTTTTATTAATACATCACTTCGAATTT 1179847  
Qy 503 CAACTTTTACATTAATATGCAAGTCCCTCAGGTG 543  
Db 1179848 TACATTTTACCATTTTATTTTAAATTTTAAATTTTAAATATG 1179888  
RESULT 27  
US-08-998-416-701  
Sequence 701, Application US/08998416  
Patent No. 6239264  
GENERAL INFORMATION:  
APPLICANT: Philippsen, Peter  
APPLICANT: Pohlmann, Rainer  
APPLICANT: Steiner, Sabine  
APPLICANT: Mohr, Christine  
APPLICANT: Wendland, Jurgen  
APPLICANT: Knechtle, Philipp  
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPIL  
TITLE OF INVENTION: AND USES THEREOF  
NUMBER OF SEQUENCES: 1152  
CORRESPONDENCE ADDRESS:  
ADDRESSER: No. 6239264artis Corporation  
STREET: 3054 Cornwallis Road  
CITY: Research Triangle Park  
STATE: No. 6239264th Carolina  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/998, 416  
FILING DATE: 24-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: CH 0016/97  
FILING DATE: 31-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38, 241  
REFERENCE/DOCKET NUMBER: PF/5-30306/N/CGC1976  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8587  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 701:  
SEQUENCE CHARACTERISTICS:

LENGTH: 701 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: PAG1462RP  
US-08-998-416-701

Query Match 6.1%; Score 37.8; DB 3; Length 701;  
Best Local Similarity 48.0%; Pred. No. 0.89;  
Matches 108; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 157 CGTCATTTGGAACAATTTTCTACTATTTTTCAGACATCTTCGGTTGAGTTCA 216  
DB 475 CTTTAAATTGATATTAAGTACTACTAAATTTTACTATAATATATTTAAGAAATCT 534  
QY 217 GCACAGTCAGTAATGCTTATTTATTTATGCTGCTCAAGAGATAGCTATTTATGATG 276  
DB 535 TAAATCTAATATTTATTTATCTAAAGTATATTAATTAATTAACCTTTTATTTATTTT 594  
QY 277 CCAATATTACTTTTATTTGCTGTTACGGCATGATTCATTAACGTACAGTACAGTTGCTCA 336  
DB 595 AATATTATTTATTTAATTAATTAATTTATTTATTTATTTAATTAATTTTGGATTA 654  
QY 337 TTATGACTACAAAGTTTAAATTTAATGAGTAATTTAAT 381  
DB 655 TAAATATCATTTAATTAATGTAATTTATTAATATTAATCTTAAT 699

RESULT 28  
US-09-134-001C-595  
Sequence 595, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 595  
LENGTH: 1401  
TYPE: DNA  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-595

Query Match 6.1%; Score 37.8; DB 4; Length 1401;  
Best Local Similarity 48.0%; Pred. No. 1.1;  
Matches 108; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 280 ATATTACTTTTATTTGCTGTTACGCAATTCATTACGTAACAAGTTGGCATTA 339  
DB 841 ATACAGCTATCAATCGCTTTTGTAGTCATTTCTTATTTACTACCTTGGTGAAGCACTA 900  
QY 340 TGACACTCAAAAGTTTTTAAATTTAAATGAGTAATTTAATTAATTAATTAAGGTTGAA 399  
DB 901 TTTTGTGAATAAAGCTGAAGATTTGGGTGTTTATGATCTTTATTTGGCTTAAA 960  
QY 400 GGAATGTTGCTTCAACAACAACCTGTTGACTTTTACGTAGTTAGTTCAAGTAAT 459  
DB 961 ACAAACTGCTTTAGGTGACACCTGGCCGTATTTAGAGTACTTTATTTGCTGGTCC 1020  
QY 460 GTTATAGTATTTGCTCATATATGATTTAAAGAGAGGTTTCA 504  
DB 1021 CTTTACCTTCTGGTCAAAATTCACACTATACAGGAAGTTAGCA 1065

RESULT 29  
US-09-107-532A-905/c  
Sequence 905, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESS: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354

COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:

NAME: Arinello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 905:

SEQUENCE CHARACTERISTICS:  
LENGTH: 2142 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:

NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...2142  
SEQUENCE DESCRIPTION: SEQ ID NO: 905:  
US-09-107-532A-905

Query Match 6.1%; Score 37.8; DB 4; Length 2142;  
Best Local Similarity 47.5%; Pred. No. 1.2;  
Matches 144; Conservative 0; Mismatches 157; Indels 2; Gaps 1;

QY 230 ATGCTTATTTATTTATGCTGCTCAAGATAGCTATTTATTTATGATGCAATATTTACTTT 289  
DB 1630 ATGATTCGTTATTTTCGATTAATTAATCTTAATGAGATCTGATTTATGTAATATAT 1571  
QY 290 TTATGCTGTTAGGCAATGATTCATTAAGTACGTAAGTTGTCATTTAGCACTCAA 349  
DB 1570 CTTTGGCTTTCATTAATAGATTTAAATTAATCTTTTGGTTATTTTAAACAATA 1511  
QY 350 AGTTTTTAATTTAAATGAGTAATTTAATTAATTAATTAATTAATTAATTAATTAATTA 409  
DB 1510 TTAGTTTATTTTAAATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1453  
QY 410 CTTCTCAACAACCTGTTGACTTTTACGTAGTTAGTTAGTTCAAGTAATTTGTTATAGTA 469  
DB 1452 GTCTCAAGTCATAAAGTATGCTCTTAATTAATTTTTCATCATCTTTTGCATATTC 1393

OY 470 TTGTTCCATATATGATTTAAAGAGAGTTTCAACTTTTACCATTTATATATGCAA 529  
DB 1392 TTTTATTTTGTAGTACTAGATGTAGTGTAGTGCATCAGAGTTTATTAATTTAA 1333  
OY 530 GTC 532  
DB 1332 GTC 1330

## RESULT 30

US-09-489-847-67/c  
Sequence 67, Application US/09489847  
Patent No. 6476195  
GENERAL INFORMATION:  
APPLICANT: Rosen et al  
TITLE OF INVENTION: 98 Human Secreted Proteins  
FILE REFERENCE: P2031P1  
CURRENT APPLICATION NUMBER: US/09/489,847  
EARLIER FILING DATE: 2000-01-24  
EARLIER APPLICATION NUMBER: PCT/US99/17130  
EARLIER FILING DATE: 1999-07-29  
EARLIER APPLICATION NUMBER: 60/094,657  
EARLIER FILING DATE: 1998-07-30  
EARLIER APPLICATION NUMBER: 60/095,486  
EARLIER FILING DATE: 1998-08-05  
EARLIER APPLICATION NUMBER: 60/096,319  
EARLIER FILING DATE: 1998-08-12  
EARLIER APPLICATION NUMBER: 60/095,454  
EARLIER FILING DATE: 1998-08-06  
EARLIER APPLICATION NUMBER: 60/095,455  
EARLIER FILING DATE: 1998-08-06  
NUMBER OF SEQ ID NOS: 376  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 67  
LENGTH: 2434  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (10)  
OTHER INFORMATION: n equals a,t,g, or c  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (27)  
OTHER INFORMATION: n equals a,t,g, or c  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (12)  
OTHER INFORMATION: n equals a,t,g, or c  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (73)  
OTHER INFORMATION: n equals a,t,g, or c  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (27)  
OTHER INFORMATION: n equals a,t,g, or c  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (75)  
OTHER INFORMATION: n equals a,t,g, or c  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (103)  
OTHER INFORMATION: n equals a,t,g, or c  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (130)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-489-847-67

Query Match 6.1%; Score 37.8; DB 4; Length 2434;  
Best Local Similarity 45.5%; Pred. No. 1.3; 162; Indels 0; Gaps 0;  
Matches 135; Conservative 0; Mismatches

OY 227 TATATGCTTATTTATTTATTTGTCACAGATAGCTATTTTATTGATGCCAATATTTAC 286

DB 2421 TT 2362  
OY 287 TTTTATGCTGTAGGCANAGATTTCATTACGTGACAGTTGGCATTTAGACTA 346  
DB 2361 TTTTTTTTTTTTTTTATGTCGTGATTTTAAATTTCTTACGGGTCATTTTACTTA 2302  
OY 347 CAAAGTTTTTAAATTTAAATGAGATTAATTTATATTAATTAAGGGTTGAAGGAATGT 406  
DB 2301 AAAAAATATTAATTAATTTTAAATGATCTTAGAAAAAATTAGCAAAAAATTGAGAAAGTT 2242  
OY 407 TTGCTTCAACAAACCTGTTGACTTTTACGTAGTTAGTTCAGTAATGTTATAG 466  
DB 2241 GAGCATATTCATGTGCAATGTTTTCATTCGAAATGTTTAAATCATTAATTTGGTT 2182  
OY 467 GTATTGTCATATATATATTAAGAGAGTTTCAACTTTTACCATTTATAT 523  
DB 2181 TCCCTCAACATACACATCAAAAATACATTTCAAAAACCTTTTTCATATATAT 2125

## RESULT 31

US-07-641-143B-1  
Sequence 1, Application US/07641143B  
Patent No. 5436000  
GENERAL INFORMATION:  
APPLICANT: Bardour, Alan G.  
APPLICANT: Bundoc, Virgilio  
TITLE OF INVENTION: Flagella-less Borrelia  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TX  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/641,143B  
FILING DATE: 11-JAN-1991  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Patterson, Melinda  
REGISTRATION NUMBER: 33,062  
REFERENCE/DOCKET NUMBER: UTSK092  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713-787-1592  
TELEFAX: 713-789-2679  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1507 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Borrelia burgdorferi  
STRAIN: HB19  
IMMEDIATE SOURCE:  
LIBRARY: PAC1  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 207..1217  
OTHER INFORMATION: /product= "Flagellin"  
FEATURE:  
NAME/KEY: RBS  
LOCATION: 194..198  
FEATURE:  
NAME/KEY: 10\_signal  
LOCATION: 146..151

FEATURE:  
NAME/KEY: -35\_signal  
LOCATION: 122..125  
US-07-641-143B-1

Query Match 6.0%; Score 37.6; DB 1; Length 1507;  
Best Local Similarity 46.8%; Pred. No. 1.2;  
Matches 118; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 217 GCACAGTCAGTATAGCTTTTATTATTATGTTGCTCAAGAGATAGCTATTATTGATG 276  
DB 1083 GCTCAATTAAGAAGTGTCTACATGACATGAGTTGTAGACGACAACTAATAGTATT 1142  
QY 277 CCAATATTACTTTTATTGCTGTACGCGATGATTCATTACGTTACAAAGTGTGCA 336  
DB 1143 TTAACACAACTGCAATGCGCATGATTCGACGCTATCAAGTCCCAATATGTTTG 1202  
QY 337 TTATGACTACAAAGCTTTTAAATTAAATGAGTAAATTAAATATATATAAAGGTTG 396  
DB 1203 TCATGCTTATGATTAATAATTCTAGTTTAAATGTTAATTAATTAACAAAGATCCCTT 1262  
QY 397 AAAGAAATGTTGCTTCTCAACAACACTTGTGACTTTAGCTTAGTTCAGTA 456  
DB 1263 AAAGATCTTTGTTTATTATTTCAGATCGCAAAATTAATTAATTTAGTATAATT 1322  
QY 457 ATTGTTATAGT 468  
DB 1323 ATAATGATGTGT 1334

## RESULT 32

US-08-124-290-1  
Sequence 1, Application US/08124290  
Patent No. 5585102

## GENERAL INFORMATION:

APPLICANT: Barbour, Alan G.  
TITLE OF INVENTION: Flagella-less Borrelia  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TX  
COUNTRY: USA  
ZIP: 77210

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/124,290  
FILING DATE:

CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/641,143  
FILING DATE: 11-JAN-1991

## CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:  
NAME: Patterson, Melinda  
REGISTRATION NUMBER: 33,062  
REFERENCE/DOCKET NUMBER: UTSK092  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713-787-1592  
TELEFAX: 713-789-2679

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1507 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:  
ORGANISM: Borrelia burgdorferi  
STRAIN: HB19  
IMMEDIATE SOURCE:  
LIBRARY: PACBI  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 207..1217  
OTHER INFORMATION: /product="Flagellin"

FEATURE:  
NAME/KEY: RBS  
LOCATION: 194..198

FEATURE:  
NAME/KEY: 10\_signal  
LOCATION: 146..151

FEATURE:  
NAME/KEY: -35\_signal  
LOCATION: 122..125  
US-08-124-290-1

Query Match 6.0%; Score 37.6; DB 1; Length 1507;  
Best Local Similarity 46.8%; Pred. No. 1.2;  
Matches 118; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 217 GCACAGTCAGTATAGCTTTTATTATTATGTTGCTCAAGAGATAGCTATTATTGATG 276  
DB 1083 GCTCAATTAAGAAGTGTCTACATGACATGAGTTGTACGACGACAACTAATAGTATT 1142  
QY 277 CCAATATTACTTTTATTGCTGTACGCGATGATTCATTACGTTACAAAGTGTGCA 336  
DB 1143 TTAACACAACTGCAATGCGCATGATTCGACGCTATCAAGTCCCAATATGTTTG 1202  
QY 337 TTATGACTACAAAGCTTTTAAATTAAATGAGTAAATTAAATATAAAGGTTG 396  
DB 1203 TCATGCTTATGATTAATAATTCTAGTTTAAATGTTAATTAATTAACAAAGATCCCTT 1262  
QY 397 AAAGAAATGTTGCTTCTCAACAACACTTGTGACTTTAGCTTAGTTCAGTA 456  
DB 1263 AAAGATCTTTGTTTATTATTTCAGATCGCAAAATTAATTAATTTAGTATAATT 1322  
QY 457 ATTGTTATAGT 468  
DB 1323 ATAATGATGTGT 1334

## RESULT 33

US-08-696-372A-1  
Sequence 1, Application US/08696372A  
Patent No. 6077515

## GENERAL INFORMATION:

APPLICANT: Barbour, Alan G.  
TITLE OF INVENTION: Flagella-less Borrelia  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: PO Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/696,372A  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/641,143  
FILING DATE: 11-JAN-1991

ATTORNEY/AGENT INFORMATION:  
NAME: Patterson, Melinda  
REGISTRATION NUMBER: 33,062  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713-787-1592  
TELEFAX: 713-787-2679  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1507 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Borrelia burgdorferi  
STRAIN: HB19  
IMMEDIATE SOURCE:  
LIBRARY: PAC1  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 207..1217  
OTHER INFORMATION: /product= "Flagellin"  
FEATURE:  
NAME/KEY: RBS  
LOCATION: 194..198  
FEATURE:  
NAME/KEY: -10\_signal  
LOCATION: 146..151  
FEATURE:  
NAME/KEY: -35\_signal  
LOCATION: 122..125  
US-08-696-372A-1

Query Match  
Best Local Similarity 6.0%; Score 37.6; DB 3; Length 1507;  
Matches 118; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 217 GCACAGTCAGTATAGCTTATTATTATTATTATTGTTCTCAAGATAGCTATTATTATTGATG 276  
DB 1083 GCTCAAAATGAAAGATGCTACATGACAGATGAGTTGTAGCGACAACCACTAATGATATT 1142  
QY 277 CCAATATTCTTTTATTGCTGTACGGCATGATTCATTTCGTGTACAAAGTGTGCA 336  
DB 1143 TTAAACCAATCTGCATGCGAATGATGCGAGGCTAATCAAGTTCCCAATATGTTTG 1202  
QY 337 TTATGACTACAAAGGTTTAAATTAAATGAGTAAATTAATTAATTAATAAAGGTTG 396  
DB 1203 TCATTCCTAGATTAATTAATCTAGTTAATGGTAAATGTAATTAACAAGATCTTT 1262  
QY 397 AAAGGATGTTGCTTCTCAACAACACTGTTGCACTTTTACGTAGTTTACGTTCAAGTA 456  
DB 1263 AAAGATCTTTTGTGTTTATTTCAGATCGCAAAATTTAATTAATTTAGATTAATT 1322  
QY 457 ATTGTTATAGT 468  
DB 1323 ATATGATGTGT 1334

RESULT 34  
US-09-107-532A-3402  
Sequence 3402; Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
City: Waltham  
STATE: Massachusetts  
COUNTRY: USA

ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Arinello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 3402:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1296 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1..1296  
SEQUENCE DESCRIPTION: SEQ ID NO: 3402:  
US-09-107-532A-3402

Query Match  
Best Local Similarity 6.0%; Score 37.4; DB 4; Length 1296;  
Matches 149; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

QY 132 TGCATTAATTCAGCGTAATGAGCGATCTTTGAAACAATTTCTACTAATATTTTAC 191  
DB 201 TGGTAATATTTTAATGTAATCAATTAAGATATTGAATAAAGAAATTGAAGA 260  
QY 192 AGAATCATTCGTTGAGTTACAGACAGCTAGTATATGCTTATTATTATGTTGC 251  
DB 261 AAAAATTAATTCAAATGATGATTAGCAAAATTAAGATGCTAAAGTAGTTAGCGAGA 320  
QY 252 TCAAGATAGCTATTATTATGATGCCAATATTCCTTTTATGCTGTACGGCATGAT 311  
DB 321 ACAAAGAAATTAATAGAGAATTTATCTGACGAATTTAAGAAATGTTTACTGGAAGC 380  
QY 312 TTCATTACGCTGACAGTTGGTGCATTATGACTACAAAGTTTAAATTAATGAG 371  
DB 381 TACAATTAATCTAAAGAAATATTTTATTAGTAGTGAGATATTACAAATTAAGAGG 440  
QY 372 TAAATTAATATATAAAGGTTGAAGAAGATGTTGCTTCTCAACAACACTGTTGC 431  
DB 441 TAAAATATAGAGAAATGATTAAGAAATTCATATTTGTCATGAAGAAATTTGCTAACA 500  
QY 432 ACTTTAGGATGTTAGTTCAAGTAATGTTATAG 466  
DB 501 AAACAATCTAAATTAAGGATGAGGATGATCTTG 535

RESULT 35  
US-08-286-325A-1  
Sequence 1; Application US/08286325A  
Patent No. 5658770  
GENERAL INFORMATION:  
APPLICANT: PREVOTIS, Fabien

```
APPLICANT: REMY, Elisabeth
APPLICANT: RITZENTHALER, Paul
TITLE OF INVENTION: Nucleic acid sequence and plasmids
TITLE OF INVENTION: comprising at least one phage resistance mechanism,
TITLE OF INVENTION: bacteria in which they are present, and their use
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: BACON & THOMAS
STREET: 625 Slaters Lane - Fourth Floor
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286.325A
FILING DATE: 04-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93 09 777
FILING DATE: 09-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: FICHTER, Richard E
REGISTRATION NUMBER: 26,382
REFERENCE/DOCKET NUMBER: REF/BDL/
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 683-0500
TELEFAX: (703) 683-1080
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1875 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Lactococcus lactis
FEATURES:
NAME/KEY: CDS
LOCATION: 202..1821
US-08-286-325A-1

Query Match          6.0%; Score 37.4; DB 1; Length 1875;
Best Local Similarity 51.5%; Pred. No. 1.5;
Matches 86; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 249 TGTCTCAAGAGTAGTATTATTGATGCCAATATTACTTTTATGCTGTACGGCATG 308
DB 75 TGTCCAGAAATGACATTTTCTGAAACCCATATTAATAATTTTGTATATCCCAATA 134
QY 309 GATTTCATTACGTGTACAAAGTTGGCATTTGACACTACCAAGGTTTAAATTAATG 368
DB 135 TATTATATATAGCTTCAATGTAAATTTTATATGATATATATAGAAATTTTAAAA 194
QY 369 GAGTAATTTATATATATAAAGGTTGAAGGAATGTTTCTCTC 415
DB 195 AAATGATGATATATATATGACTTTAAACATATGTTAAGCATATC 241

RESULT 36
US-08-286-325A-7
; Sequence 7, Application US/08286325A
; Patent No. 5658770
; GENERAL INFORMATION:
; APPLICANT: PREVOTS, Fabien
; APPLICANT: REMY, Elisabeth
; APPLICANT: RITZENTHALER, Paul
```

```
TITLE OF INVENTION: Nucleic acid sequence and plasmids
TITLE OF INVENTION: comprising at least one phage resistance mechanism,
TITLE OF INVENTION: bacteria in which they are present, and their use
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: BACON & THOMAS
STREET: 625 Slaters Lane - Fourth Floor
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286.325A
FILING DATE: 04-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93 09 777
FILING DATE: 09-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: FICHTER, Richard E
REGISTRATION NUMBER: 26,382
REFERENCE/DOCKET NUMBER: REF/BDL/
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 683-0500
TELEFAX: (703) 683-1080
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3234 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Lactococcus lactis
US-08-286-325A-7

Query Match          6.0%; Score 37.4; DB 1; Length 3234;
Best Local Similarity 51.5%; Pred. No. 1.7;
Matches 86; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 249 TGTCTCAAGAGTAGTATTATTGATGCCAATATTACTTTTATGCTGTACGGCATG 308
DB 587 TGTCCAGAAATGACATTTTCTGAAACCCATATTAATAATTTTGTATATCCCAATA 646
QY 309 GATTTCATTACGTGTACAAAGTTGGCATTTGACACTACCAAGGTTTAAATTAATG 368
DB 647 TATTATATATAGCTTCAATGTAAATTTTATATGATATATATAGAAATTTTAAAA 706
QY 369 GAGTAATTTATATATATAAAGGTTGAAGGAATGTTTCTCTC 415
DB 707 AAATGATGATATATATATGACTTTAAACATATGTTAAGCATATC 753

RESULT 37
US-09-134-001C-2508
; Sequence 2508, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT FILING DATE: 1998-08-13
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
```

PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 2508  
LENGTH: 576  
TYPE: DNA  
ORGANISM: *Staphylococcus epidermidis*  
US-09-134-001C-2508

Query Match 6.0%; Score 37.2; DB 4; Length 576;  
Best Local Similarity 44.7%; Pred. No. 1.2;  
Matches 144; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 210 GGTTCAGACAGTGTATGCTTATTTATGTCGCAAGATGCTATTT 269  
DB 87 GGTAAAGATCAACAAGACATGATTAAGAAAGTTGTTATCAGATATAGGTC 146  
QY 270 ATTGATGCCAATATTACTTTTATGCTGTACGGCAGATTCATTCGTGCAAGT 329  
DB 147 AATTATGACAGATTTTTCCTAATGATGAGCAACCTTAATTTGACCCATGCTTT 206  
QY 330 TGGTCATTATGCACTACCAAGCTTTTAATTTAAATGAGTAATTTAATATATAA 389  
DB 207 AGCGATGATAAGAAATTTAGTTTCAATATAAAAGCTAGTTTCCCTTTATATAGCTTG 266  
QY 390 AGGTTGAAGAAGATGTTGCTTCAACAACCTGTGCACTTTACGATTTAGT 449  
DB 267 TGGGTTATCGATATATCCACATCATTTACCTGTTGTAAGTAATTTGTTAACTTGT 326  
QY 450 TCAAGTAATTTGTTATAGTATGTTCCATATATGATTATATAAGAGAGCTTTCAACTT 509  
DB 327 ATCAGCAGATTACTTTGTTGATTTGAATTTGTAATTTTAAAGCGATGTTATACCTTA 386  
QY 510 TTTACCATTAATATATGCAAGT 531  
DB 387 TTTATTTCTCTATTAGCTAGT 408

## RESULT 38

US-08-213-419B-3/c  
Sequence 3, Application US/08213419B  
Patent No. 6333406  
GENERAL INFORMATION:  
APPLICANT: Inselburg, J. et al.  
TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM  
FILE REFERENCE: J11-002CNP  
CURRENT APPLICATION NUMBER: US/08/213,419B  
CURRENT FILING DATE: 1994-03-14  
PRIOR APPLICATION NUMBER: US 07/870,506  
PRIOR FILING DATE: 1992-04-17  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 6124  
TYPE: DNA  
ORGANISM: *Plasmodium falciparum*  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (2407)..(2439)  
NAME/KEY: CDS  
LOCATION: (2598)..(3404)  
NAME/KEY: CDS  
LOCATION: (3580)..(3720)  
NAME/KEY: CDS  
LOCATION: (3850)..(5835)  
US-08-213-419B-3

Query Match 6.0%; Score 37.2; DB 4; Length 6124;  
Best Local Similarity 45.3%; Pred. No. 2.3;  
Matches 135; Conservative 0; Mismatches 163; Indels 0; Gaps 0;  
QY 144 AGCGTAATGAGACGCTATTGAAACAATTTCTACTATATTTTACGAATCATTTGG 203

DB 2425 AGAAATATATGACTTCTATTTTGAATATGAAATTAAGTATATAATTTTAAAGACT 2366  
QY 204 GTTAGAGTTACAGACAGTATATGCTTATTTATATGTCGAAGATAGC 263  
DB 2365 TTAATTTTGAATATGCTGATTAATTTTCTTATATGTAATATATATATAT 2306  
QY 264 TATTTATTTGATGCAATATTACTTTTATTTGCTGTACGGCAGATTCATTCATGCT 323  
DB 2305 TTAATATATATATATATTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2246  
QY 324 ACAAGTTGTCATTATGACTACCAAGCTTTTAAATTTAAATGAGTAATTAAT 383  
DB 2245 TAAATTAATAAATTTATATATATATATATATATATATATATATATATATAT 2186  
QY 384 AATAAAGGTTGAAGAAGATGTTGCTTCAACAACCTGTGCACTTTAGT 441  
DB 2185 TTTAATTTATTTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2128

## RESULT 39

US-09-599-360B-7  
Sequence 7, Application US/09599360B  
Patent No. 6548633  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Bouguerelet, L.  
APPLICANT: Jobert, S.  
TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides  
FILE REFERENCE: GENSET.050CP3  
CURRENT APPLICATION NUMBER: US/09/599,360B  
CURRENT FILING DATE: 2000-06-21  
PRIOR APPLICATION NUMBER: 60/113,686  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: 60/141,032  
PRIOR FILING DATE: 1999-06-25  
PRIOR APPLICATION NUMBER: 09/469,099  
PRIOR FILING DATE: 1999-12-21  
NUMBER OF SEQ ID NOS: 123  
SOFTWARE: Patent.pm  
SEQ ID NO 7  
LENGTH: 1918  
TYPE: DNA  
ORGANISM: *Homo sapiens*  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 238..612  
NAME/KEY: sig\_peptide  
LOCATION: 238..348  
OTHER INFORMATION: Von Heijne matrix  
OTHER INFORMATION: score 9.4  
NAME/KEY: polyA\_signal  
LOCATION: 1885..1890  
NAME/KEY: polyA\_site  
LOCATION: 1905..1918  
NAME/KEY: misc\_feature  
LOCATION: 945..1624  
OTHER INFORMATION: n=a, g, c or t  
US-09-599-360B-7

Query Match 5.9%; Score 37; DB 4; Length 1918;  
Best Local Similarity 50.3%; Pred. No. 1.9;  
Matches 91; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 338 TATGACTACAAAGCTTTTAAATTTAAATGAGTAATTTAATATAATTAAGGTTGA 397  
DB 1723 TTTGTAATATCGTGCATTAATAATTAAGTGAACAACTATTTTAAAGCAGTTA 1782  
QY 388 AAGAATGTTGCTTCTCAACAACATTTTGACCTTTTACGTATTTAGTTCAAGTAA 457  
DB 1783 AATTATCTCTGATTTGTGAACATTAAGACTTTCTAGAAATTTACTTATTTCTGTACTTA 1842



QY	458	TTGTTAAAGCATTTGTTCCATATGATTTATTAAGAAGAGTTTCAACTTTTACCAT	517
Db	1843	AATTTTTCTAATGACACATTAATCTTTGTATCAGAAAATATTAATGCGATGATTTT	1902
QY	518	T 518	
Db	1903	T 1903	

```

RESULT 40
US-09-599-360B-60
; Sequence 60, Application US/09599360B
; Patent No. 6548633
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bougaletret, L.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: GENSET.050CP3
; CURRENT APPLICATION NUMBER: US/09/599,360B
; CURRENT FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: 60/113,686
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/141,032
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/469,099
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: Patent.pm
; SEQ ID NO 60
; LENGTH: 1918
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 238..612
; NAME/KEY: sig_peptide
; LOCATION: 238..348
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 9.40
; OTHER INFORMATION: seq LLCVSVASQSLSS/QD
; NAME/KEY: polyA_signal
; LOCATION: 1885..1890
; NAME/KEY: polyA_site
; LOCATION: 1905..1918
; NAME/KEY: misc_feature
; LOCATION: 945..1624
; OTHER INFORMATION: n=a, g, c or t
; US-09-599-360B-60

```

	Query Match	5.9%	Score 37	DB 4	Length 1918
	Best Local Similarity	50.3%	Pred. No. 1.9		
	Matches	91	Conservative	0	Mismatches 90
					Indels 0
					Gaps 0
QY	338	TATGACTACAAAGGTTTTAAATTTAAATGAGTAATTAATATTAATAAAGGTTGA	397		
Db	1723	TTTGTAAATACGTGTCATAAAAATTAATTAAGAGGAACAACATATTTTAAAGCAGTTA	1782		
QY	398	AAGAAATGTTGCTTCACAAACAACCTGTGCCTTTACGTGTTAGTTCAGATTA	457		
Db	1783	AATTAATCTCTATATGTGTAACCTAAGACTTTCTAGAAATTTACTTAATTAATCTGTACTTA	1842		
QY	458	TTGTTATAGTATTGTTCCATATATGATTAATAAAGAGAGTTTCAACTTTTACCAT	517		
Db	1843	AATTTTTTCTAATGAACAATATACCTTTGTAATCAGAAATATTAATTAATGATGATATTT	1902		
QY	518	T 518			
	1903	T 1903			

Search completed: October 4, 2003, 05:14:34  
Job time : 61 secs

**THIS PAGE BLANK (USPTO)**